United States Department of State

Washington, D.C. 20520

February 29, 2024

Case No. FL-2023-00013

Reed Rubinstein
America First Legal Foundation
611 Pennsylvania Avenue, SE, #231
Washington, DC 20003

Dear Mr. Rubinstein:

As we noted in our letter dated January 31, 2024, we are processing your request for material under the Freedom of Information Act ("FOIA"), 5 U.S.C. § 552. The Department of State ("Department") has identified five additional responsive records subject to the FOIA. Upon review, we have determined that one record may be released in full and four records may be released in part.

An enclosure explains the FOIA exemptions and other grounds for withholding material. Where we have made redactions, the applicable FOIA exemptions are marked on each record. Where applicable, the Department has considered the foreseeable harm standard when reviewing these records and applying FOIA exemptions. All non-exempt material that is reasonably segregable from the exempt material has been released and is enclosed.

We will keep you informed as your case progresses. If you have any questions, your attorney may contact Kevin Bell, U.S. Department of Justice Trial Attorney, at kevin.k.bell@usdoj.gov and (202) 305-8613. Please refer to the case number, FL-2023-00013, and the civil action number, 22-cv-03386, in all correspondence about this case.

Sincerely,

Diamonece Hickson
Chief, Litigation and Appeals Branch
Office of Information Programs and Services

Enclosures: As stated.

Freedom of Information Act (5 U.S.C. § 552) and Privacy Act (5 U.S.C. § 552a)

FOIA Exemptions

(b)(1) Information specifically authorized by an executive order to be kept classified in the interest of national defense or foreign policy. Executive Order 13526 includes the following classification categories:

1.4(a) Military plans, systems, or operations
1.4(b) Foreign government information
1.4(c) Intelligence activities, sources or methods, or cryptology
1.4(d) Foreign relations or foreign activities of the US, including confidential sources
1.4(e) Scientific, technological, or economic matters relating to national security, including defense against transnational terrorism
1.4(f) U.S. Government programs for safeguarding nuclear materials or facilities
1.4(g) Vulnerabilities or capabilities of systems, installations, infrastructures, projects, plans, or protection services relating to US national security, including defense against transnational terrorism
1.4(h) Weapons of mass destruction

(b)(2) Related solely to the internal personnel rules and practices of an agency

(b)(3) Specifically exempted from disclosure by statute (other than 5 USC 552), for example:

ARMSEXP	Arms Export Control Act, 50a USC 2411(c)
CIA PERS/ORG	Central Intelligence Agency Act of 1949, 50 USC 403(g)
EXPORT CONTROL	Export Administration Act of 1979, 50 USC App. Sec. 2411(c)
FS ACT	Foreign Service Act of 1980, 22 USC 4004
INA	Immigration and Nationality Act, 8 USC 1202(f), Sec. 222(f)
IRAN	Iran Claims Settlement Act, Public Law 99-99, Sec. 505

(b)(4) Trade secrets and confidential commercial or financial information

(b)(5) Interagency or intra-agency communications forming part of the deliberative process, attorney-client privilege, or attorney work product

(b)(6) Personal privacy information

(b)(7) Law enforcement information whose disclosure would:
(A) interfere with enforcement proceedings
(B) deprive a person of a fair trial
(C) constitute an unwarranted invasion of personal privacy
(D) disclose confidential sources
(E) disclose investigation techniques
(F) endanger life or physical safety of an individual

(b)(8) Prepared by or for a government agency regulating or supervising financial institutions

(b)(9) Geological and geophysical information and data, including maps, concerning wells

Other Grounds for Withholding

NR Material not responsive to a FOIA request excised with the agreement of the requester

Privacy Act Exemptions

(d)(5)	Information compiled in reasonable anticipation of a civil action or proceeding
(j)(1)	Information maintained by the CIA
(j)(2)	Enforcement of criminal law, including efforts to prevent, control, or reduce crime or apprehend criminals, except records of arrest
(k)(1)	Classified pursuant to E.O. 13526 in the interest of national defense or foreign policy
(k)(2)	Investigatory material compiled for law enforcement purposes
(k)(3)	Regarding protective services to the President of the United States or other individual pursuant to Title 18, U.S.C., Section 3056
(k)(4)	Required by statute to be maintained and used solely as statistical records
(k)(5)	Investigatory material compiled solely for the purpose of determining suitability, eligibility, or qualifications for Federal civilian employment or access to classified information, the disclosure of which would reveal the identity of the person who furnished information pursuant to a promise that his identity would be held in confidence
(k)(6)	Testing or examination material used solely to determine individual qualifications for appointment or promotion in the Federal service, the disclosure of which would compromise the testing or examination process
(k)(7)	Evaluation material used to determine potential for promotion in the armed services

From: (b)(6) @state.gov>
To: (b)(6) @state.gov>
Subject: Re: DRAFT Email for DHS? Others?
Date: Tue, 27 Oct 2020 19:03:50 +0000

And given 40% of GNews' site traffic comes from social media (80% of that from Twitter), all of this is ripe for inauthentic behavior.

From (b)(6)
Sent: Tuesday, October 27, 2020 2:54 PM
To (b)(6) @state.gov>
Subject: DRAFT Email for DHS? Others?

See below as mentioned.

NAME

(b)(5)

Best Regards,

(b)(6)

Counter Disinformation Analyst | Russia Team
Global Engagement Center

U.S. Department of State
Contractor: All Native Group

Cell (b)(6)
Personal: (b)(6)
Sender: (b)(6) @state.gov>
Recipient: (b)(6) @state.gov>

From: (b)(6) state.gov>
To: (b)(6) state.gov>
(b)(6) state.gov>;
(b)(6) state.gov>;
CC: (b)(6) state.gov>;
(b)(6) state.gov>;
(b)(6) state.gov>
Subject: Fw: FireEye Report: Alleged Russian 'NAEBC' News Site and Personas Remain Active, Continue to Promote Content Related to U.S. Election
Date: Fri, 23 Oct 2020 13:00:25 +0000

(b)(6) -- let's push this out to our election synch distro list which includes those agencies, as authorized below, if not covered separately via other I2C2 mechanisms. Thanks, (b)(6)
From: (b)(6) state.gov>
Sent: Friday, October 23, 2020 8:46 AM
To: GEC-Russia Team <GEC-RussiaTeam@state.gov>; (b)(6) state.gov>
Cc: (b)(6) state.gov> (b)(6) state.gov>; (b)(6) (b)(6) mandiant.com>; GEC_DataAnalytics <GECDataAnalytics@state.gov>; GEC I2C2 Internal <GECI2C2Internal@state.gov>
Subject: FireEye Report: Alleged Russian 'NAEBC' News Site and Personas Remain Active, Continue to Promote Content Related to U.S. Election

Good Morning -
Please see the attached FireEye finished intelligence report that was recently published to our FireEye Intelligence Portal.
The report details continued information operation activity related to the U.S. election from the inauthentic news outlet "Newsroom for American and European Based Citizens" (NAEBC) despite the outlet's public exposure in October as a site allegedly controlled by foreign actors. According to *Reuters* reporting on an alleged FBI investigation, the outlet is run by individuals associated with the Russian Internet Research Agency (IRA). We have also observed inauthentic personas affiliated with NAEBC remain active on the social media platforms Gab and Parler, including by promoting articles pertaining to the unverified New York Post story about documents allegedly obtained from the laptop of former Vice President Joe Biden's son, Hunter Biden.
This report is approved to be shared in FULL with US Federal Civilian Agencies and Departments to include DHS, FBI, USAGM, USAID etc.
Approx. 1500 characters or two paragraphs of the report are approved to be shared with all other USG and FVEY partners outside of US Federal Civilian Agencies and Departments. This includes with the IC, DoD and combatant commands. If interested in this option of sharing two paragraphs more widely, I'm happy to assist with a derivative work product.
Please let me know if you have any questions about sharing or the report content. Also interested in any feedback on the report!
Thanks,
(b)(6)

(b)(6)
Global Engagement Center
U.S. Department of State
FireEye Inc.
(b)(6)

Sender: (b)(6) @state.gov>

Recipient: (b)(6) @state.gov>;
(b)(6) @state.gov>;
(b)(6) @state.gov>;
(b)(6) @state.gov>;
(b)(6) @state.gov>;
(b)(6) @state.gov>

(b)(6)

From: Ross Ewald (b)(6)

To: @state.gov>

Subject: EIP-536 Five Russian-linked or aligned proxy outlets amplify narratives surrounding Hunter Biden scandals.

Date: Tue, 17 Nov 2020 07:23:06 +0000

Ross Ewald resolved this as Out of Scope.

View request · Turn off this request's notifications

This is shared with GEC and (b)(6)

Powered by Jira Service Management

Sender: Ross Ewald (b)(6)

Recipient: (b)(6) state.gov>

CENTER OF EXCELLENCE ON DEMOCRACY, HUMAN RIGHTS, AND GOVERNANCE

DISINFORMATION PRIMER

February 2021

USAID
FROM THE AMERICAN PEOPLE

ACKNOWLEDGEMENTS

There are many people to thank in terms of their contributions to this primer. It was conceived and developed by Joshua Machleder and Shannon Maguire at USAID, together with the NORC at the University of Chicago team including Susan Abbott, Renée Hendley, and Luis Camacho.

We thank the following individuals for sharing their time, opinions, and expertise: Deepanjalie Abeywardana (Verite Research, Sri Lanka); Akintunde Akanni (Lagos State University), Daniel Arnaudo (National Democratic Institute); Manisha Aryal (Chemonics); Rafiq Copeland (Internews); Marius Dragomir (Central European University, Center for Media, Data and Society); Dejan Georgievski (Media Development Centre Skopje); Dean Jackson (National Endowment for Democracy); Rasto Kuzel (Memo 98, Slovakia); Shanthi Kalathil (National Endowment for Democracy); Gillian McCormack (Internews); Michael Mirny (IREX); Sarah Oates (University of Maryland); Igor Rozkladaj (Center for Democracy and Rule of Law, Kyiv); Bruce Sherman (USIP Peace Tech Lab); Juni Soehardjo (media lawyer, Indonesia); Tara Susman-Pena (IREX); Emeka Umejei (University of Ghana, Accra); Herman Wasserman (University of Capetown); Nancy Watzman (FirstDraft News); Bob Wekesa (University the Witwatersrand); and Tim Weninger (University of Notre Dame).

We also want to thank the United States Department of State Global Engagement Center for their review and feedback. We appreciate the feedback and advice offered by Nicholas Glavin, Mary-Beth Polley, and Phillip Tolentino.

In addition, numerous current and former staff at USAID contributed to the development of the primer. We greatly appreciate the opinions and feedback from the following: Mariam Afrasiabi, Kora Andrieu, Matthew Baker, Keti Bakradze, Jared Ford, Maher M. Frijat, Andrew Greer, Adam Kaplan, Lauren Kirby, Nadereh Lee, Taly Lind, Amy Malessa, Laura McKechnie, Michael McNulty, Kyle Novak, Diana Parzik, Marko Pjevic, Lisa Poggiali, Joseph Scheibel, Gloria Steele, Samantha Turner, Sara Werth, Thomas White, and Johanna Wilkie.

TABLE OF CONTENTS

TABLE OF FIGURES

DISINFORMATION PRIMER

This primer presents an overview of disinformation culture to give readers a sense of key concepts, terminology, select case studies, and programmatic design options. Disinformation is by no means new. Although social media platforms have emerged as the most efficient spreaders of false information, disinformation is also spread through analog media such as radio, television, and newspapers. It is, however, the combination of traditional analog media, in concert with new digital technologies, that allows information to spread faster and more broadly (even across borders) in unprecedented ways. Experts have described this phenomenon as "information disorder," a condition in which truth and facts coexist in a milieu of misinformation and disinformation—conspiracy theories, lies, propaganda, and half-truths. They have labeled its ability to undermine democracy and individual autonomy "a wicked problem," i.e., a problem that is difficult and complex, such as poverty or climate change. Despite the immensity of the challenge, there are promising ways that journalists, civil society organizations, technology specialists, and governments are finding to prevent and counter misinformation and disinformation. This primer presents several programmatic ideas to consider for standalone or integrative approaches as part of democracy and governance-related programming.

INTRODUCTION: HOW TO USE THIS PRIMER

This primer is compiled with the intention of helping USAID staff and partners to understand the basics of disinformation, how it is spread, why it is spread, and how programming can help reduce its damaging impact on societies around the world. It is organized into seven parts that each focus on illuminating the issue with insights from leading thinkers.

These insights are supplemented with resources, case studies, and examples to illustrate different dimensions of the problem and to enable readers to pursue deeper discussions and resources that can help their programs and projects. The primer and its many annexes can be used as a guide or reference, and its modular design can supplement training programs aimed at different aspects of the disinformation conundrum.

Photo: ©2018 Unsplash/Abhijith S. Nair

I. PART ONE: WHY DOES DISINFORMATION MATTER?

Part One explores how the well-worn and known tactics of disinformation are being adapted and used around the world. Evidence is mounting that "false information can reach more people, penetrate more deeply into social networks, and spread much faster than any time in history."[1]

Experts from academia, government, civil society, and media agree that disinformation is a problem with social, political, and economic ramifications. A study done by Prevency, a German international consulting company for reputational risk and crisis management, found that disinformation costs the global economy $78 billion per year, including in share price loss, brand reputation management, and investment in political disinformation campaigns.[2]

USAID staff and partners around the world need a working knowledge of the scope and form of disinformation since it impacts many levels of programming and interventions across all development sectors. While it is daunting to define terms, this primer provides key terminology and tools to better identify ways to counter and prevent it.

Disinformation comes from both homegrown and foreign sources. *Foreign Policy* noted in a recent article that "as research has increasingly shown, homegrown disinformation is making democracy sicker than any foreign efforts can."[3] The article goes on to point out:

> "There are immense incentives for disinformation built into democratic institutions themselves. Treating disinformation as an alien disease ignores the fact that it is perfectly compatible with democratic norms and thrives inside democratic states. A recent report[4] by the Oxford Internet Institute, for example, found that politicians inside 45 democracies have used social media for 'amassing fake followers or spreading manipulated media to garner voter support.'[5]"

Disinformation is a core challenge for democracy, rights, and governance promotion, yet it is not the only problem. Other key information challenges include censorship and freedom of expression; internet freedom and digital rights (including throttling and internet shutdowns); political polarization; and the demise of traditional journalism business models and related new challenges of the financial viability of the news industry in the digital age. Each of these challenges creates fertile ground for, or amplifies disinformation by, limiting the free and open access to facts, data, and information in our societies.

As the spread of disinformation online has grown rapidly in recent years, global internet freedom has been declining rapidly (for the ninth year in a row in 2019). Since 2016, Freedom House has reported on new governments contributing to the spread of disinformation. Freedom House also has observed new malicious actors taking advantage of the failure of democratic states to successfully regulate online campaign finance and transparency rules that are essential for democratic elections. This trend is worrisome. Many

more democratic leaders are employing this strategy domestically. In this way, democratic governments are also falling prey to state-sponsored disinformation because they cannot use more draconian methods to exercise power.

Additionally, repressive governments have gained access to new tools to collect and track data on entire population sets and are utilizing them to effectively increase popular support for themselves. They use social media surveillance tools and artificial intelligence to "identify perceived threats and silence undesirable expression." This trend has created an environment where civil rights are being abused and activists are being repressed and denied the possibilities of the digital sphere for a variety of religious, social, and political speech. Of the 65 countries Freedom House assessed in 2019 in its *Freedom on the Net Report*, 47 (the highest number to date) have arrested online users for religious, social, or political posts. Such cooptation and control of the digital space further allows authoritarian leaders to engage in domestic disinformation campaigns more easily and widely.

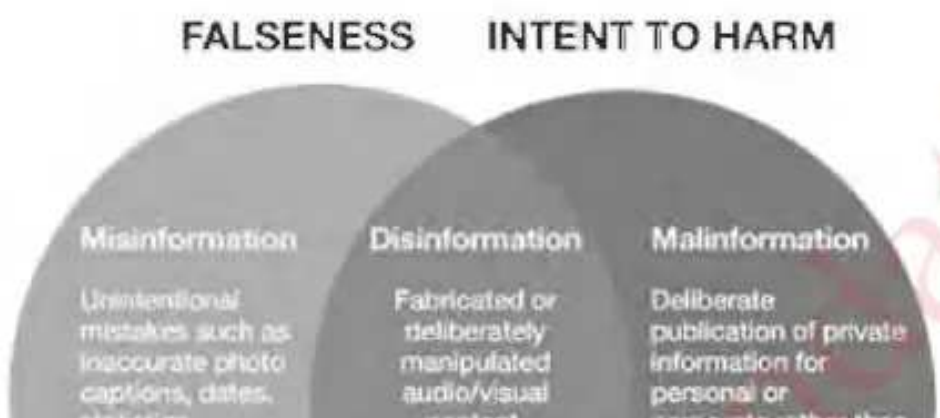
A. DISINFORMATION, MISINFORMATION, AND MALINFORMATION

Disinformation, or **information that is shared with the intent to mislead people**, is increasingly a global phenomenon. It has become more prevalent with the rise of social media and the digital economy and a lack of digital and media literacy among consumers of online media.[6] Disinformation is often used as a catch-all term for all false information, but it is distinguished from misinformation by its purposeful intent to deceive. **Misinformation**, on the other hand, is **false information spread by someone who believes false information to be true**. (See Figure 1.) The impact of disinformation and misinformation can be the same.[7] Whether false information is shared intentionally or not, it is still dangerous.

Additionally, **"malinformation"** is deliberate publication of private information for personal or private interest, as well as the deliberate manipulation of genuine content. This is often done by moving private or revealing information about an individual, taken out of context, into the public sphere.

Researchers across

Figure 1: Information disorder



Source: https://internews.org/impact/disinformation

Researchers across disciplines have shown in a variety of ways how networked disinformation capitalizes on predictable human behavior in digital spaces. In a Massachusetts Institute of Technology (MIT) study, **researchers found that false information travels on average six times faster than authentic, true new stories**. The study's data showed false information "diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information," suggesting people are more likely to share false stories for their novelty and because they "inspire fear, disgust, and surprise in replies." **Media literacy**, or **the ability to methodically consider the meaning and source of a post or news article**, is the most important factor in identifying false or misleading news. As a Yale University study found, "Susceptibility to fake news [false news] is driven more by lazy thinking than it is by partisan bias *per se*."[8]

B. Disinformation and its impact on democracy, rights, and governance

Posing an existential threat to democracy, the spread of misinformation and disinformation is having an "absolutely corrosive" impact on the institutions and norms that enable democratic governance.[9] The impact it has on prospects for democratic development is an urgent matter for those involved in strengthening democratic institutions. Ideally, in a healthy democracy, citizens and policymakers can draw from a common set of facts to deliberate and make decisions. In an era where mis- and disinformation is so prevalent, democratic progress and order is threatened by faulty information—conspiracies, lies, half-truths, distortions, and propaganda.

> "We need to be as cautious about our information hygiene as we are about our hand hygiene, for the sake of public health and democracy."
>
> —Joyce Fegan, *Irish Examiner*, September 19, 2020

Most worryingly, disinformation is a significant force that can undermine democracy and good governance, free and fair elections, access to information, rule of law, protection of human rights, independent media, and civil society action. Critical to every aspect of good

governance, information integrity enables political parties to debate and share ideas, concerns, and solutions. It opens opportunities for citizens to influence public policy dialogue. It promotes economic innovation as entrepreneurs refine and improve on the goods we produce. And it enables governments to respond effectively to public health and other emergencies.

Four ways in which disinformation impacts democratic development

- Interferes directly with the ability of democratic societies to determine what is in the public interest through open discussion
- Leads to a loss of information

integrity and often to impediments of press freedom
- Can interfere with civic society and distort political engagement
- Exacerbates polarization and social fracturing

Democratic societies rely on journalists, media outlets, and bloggers to help shape local and national dialogue, shine a light on corruption, and provide truthful, accurate information that can inform people and help them make the decisions needed to live and thrive. Yet, the public sphere has taken on an increasingly toxic and polarized quality. The nature of how people access information is changing along with the information technology boom and the decline of traditional print media. Because traditional information systems are failing, some opinion leaders are casting doubt on media, which, in turn, impacts USAID programming and funding choices.

Our technology-enabled society, with all the vaunted benefits of increased connection, efficiency, and transparency, has also led to erosion of individual privacy, trolling, cyberbullying, cyber or information warfare, and dubious and deceitful abuse and misuse of tech platforms in deliberate efforts to erode democracy and public trust and extort money and power. As numerous scholars have argued, tech platforms have preferred profit over their users, failing to provide even basic controls to help support civic engagement over extremist speech. Indeed, studies such as Siva Vaidhyanathan's *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy* demonstrate that social media platforms find extremism far more engaging—and hence more profitable—so their platform design encourages it.[10]

Today's digital communications and media landscape is complex and has given rise to a new set of challenges and considerations for democracy support. Across all USAID programming countries, this requires a solid understanding of disinformation and robust approaches for countering and preventing it.

Two current examples from Brazil and Myanmar illustrate the threat posed to democracy by rampant disinformation. The way in which disinformation is manifest is very much context based. One must consider how information is consumed in a particular country or community. For example, in Myanmar effective information manipulation relies on Facebook, which is almost synonymous with the internet since it is provided free with a mobile phone data plan. Recent developments with the Myanmar military displacing civilian leadership in a coup in January 2021 have underlined issues of internet freedom and disinformation. Likewise, in Brazil, cheap data plans that only include access to the likes of WhatsApp and Facebook makes citizens more likely to consume social media in which false claims are echoed by unreliable sources.

Disinformation impacts the prospects for democratic development in a number of ways:

1. **It undermines trust in democratic institutions by reducing their credibility and legitimacy in the eyes of the public.** Disinformation interferes directly with the ability of democratic societies to determine what is in the public interest by dominating and distorting the public discourse and corrupting the process of democratic decision-making.

The spread of disinformation is a tactic that authoritarians use to dominate people and societies. [11] Governments also sometimes use it as a tool of foreign policy.[12] When the strategy leads to political success, it provides motivation for the winners to restrict the

free flow of information by those who would dispute them and to undermine constitutional protections for free speech.[13]

3. **It leads to a loss of information integrity.** Online news platforms have disrupted the traditional media landscape. Government officials and journalists are not the sole information gatekeepers anymore. As such, citizens require a new level of information or media literacy to evaluate the veracity of claims made on the internet. False beliefs spread across the internet because almost anything is being promoted by someone. Authoritarian leaders add to the loss of information integrity by delegitimizing the media, claiming news sources are faulty or corrupt, i.e., the weaponization of "fake news." The loss of information integrity itself further undermines trust in the media's ability to provide fact-based information. **It leads to a loss of press freedom.** The weaponization of "fake news" (calling a partisan or otherwise opinion-heavy article or outlet "fake news" in order to discredit it) has also led some governments to propose or pass anti "fake news" bills, which have had a chilling effect on freedom of speech and are used to target or silence independent media.

4. **It can distort political and civic engagement.** Social media platforms offer democratic benefits by connecting citizens with each other in ways more easily possible in a digital space, encouraging voter turnout, and giving voice to minority viewpoints. However, in conjunction with disinformation, the same platforms can provide the means for suppression of civic and political engagement. The use of trolls, doxing, flooding, and other tactics have resulted in a dramatic reduction in constructive social and political engagement. Simple widespread mistrust about the accuracy and authenticity of online information may be enough to demotivate political engagement.[14]

5. **It exacerbates polarization and social fracturing.** Information technology creates many opportunities to engage with and learn from different perspectives. On the other hand, new information technology has been used to reinforce stereotypes and create insular communities with similar values, histories, and experiences, providing a home for disaffected populations to promote their own views. This is further complicated by "filter bubbles" and "echo chambers"[15] created by social media, in which false claims are repeated and magnified, increasing polarization and making democratic discourse more difficult as citizens turn to different sets of false information as facts.

6. **It can have a disproportionate impact on marginalized populations, resulting in online violence, intimidation, and harassment using false narratives.** Disinformation on social media often involves targeted harassment campaigns that seek to silence and marginalize opposing opinions and/or specific groups in society, such as women or ethnic groups, and make it appear that disinformation actors preside over greater consensus. Such harassment and attempts to silence voices have been used to discourage and discredit women candidates for political office in many countries.

PART TWO: UNDERSTANDING INFORMATION DISORDER

Although disinformation has become a hot button issue over the past several years, the manipulation of media has been used as a political tactic in the past and is by no means a new phenomenon. Information disorder, a term coined by Claire Wardle[16] an expert on misinformation and disinformation and co-founder of First Draft, refers to the current media climate and the ways in which the media ecosystem is polluted.[17]

Photo: ©2017 Unsplash/Kayla Velasquez

This term offers an alternative to the term "fake news," which has been coined and promoted for political purposes. As noted by *The Conversation*, "Not only do different people have opposing views about the meaning of "fake news," in practice the term undermines the intellectual values of democracy and there is a real possibility that it means nothing. We would be better off if we stopped using it."[18] Furthermore, as noted by assistant director-general of communication and information at UNESCO Frank La Rue:

> "Fake news is a bad term primarily because it is a trap. It is not news. Just the term generates mistrust of the press and of the work of journalists. Political leaders have started using the term against the press, which is especially serious. This is a crucial moment when we have to defend journalism. We have to promote a journalism of honesty, a journalism that is seen to build the truth."[19]

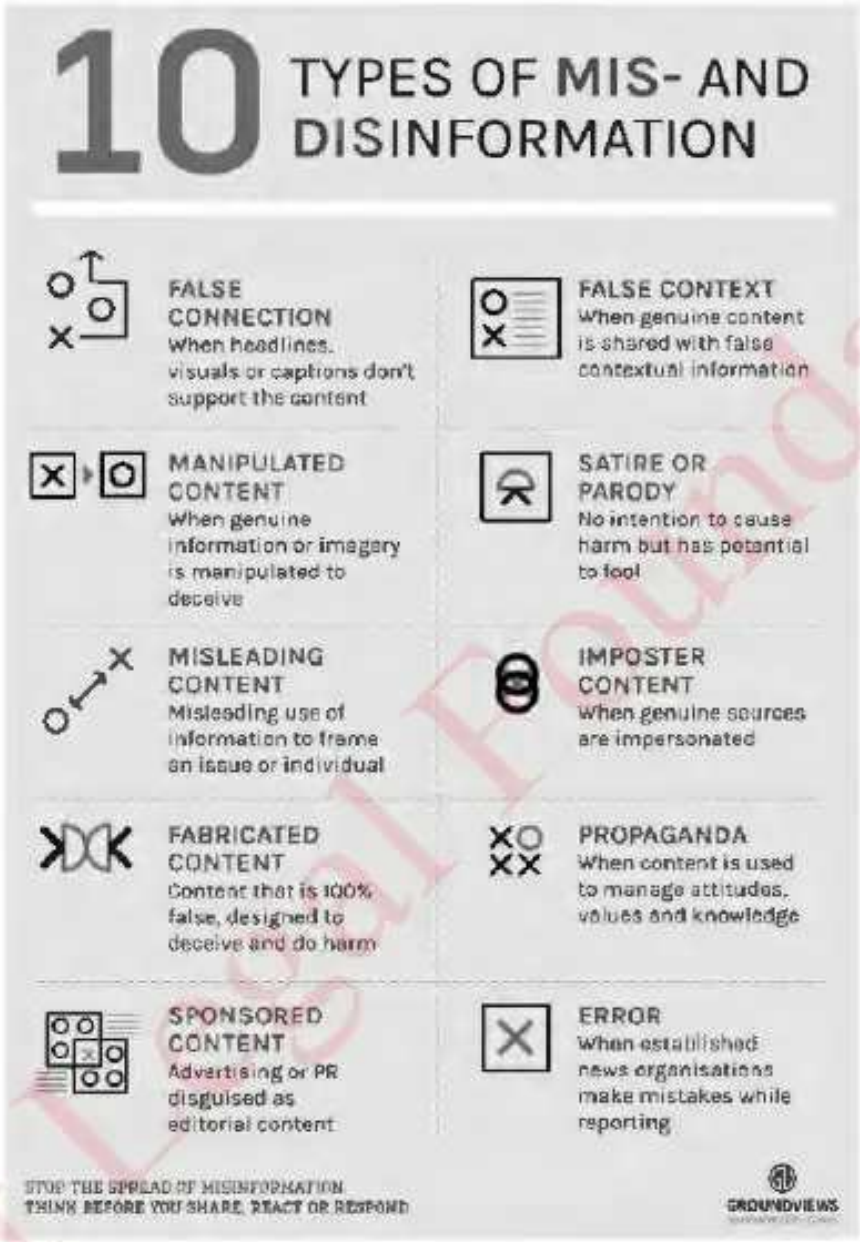
Information disorder. A condition in which truth and facts coexist in a milieu of misinformation and disinformation—conspiracy theories, lies, propaganda, and half-truths. In fact, Groundviews identified 10 types of misinformation and disinformation. (See Figure 2. First Draft News also prepared illustrative examples of each type; see Annex 2: Types of Misinformation & Disinformation, to study them.)

The threats of information disorder have worsened as social media and internet use become more ubiquitous and as digital technology writ large has taken on a bigger role in democracy and governance programming. It is a central area of study to understand how and why there has been such an expansive erosion of democracy over the past 10 years[20]—since 2010,

the number of "Free" nations in Freedom House's Freedom in the World index has decreased by six percent while "Not Free" nations increased by 10 percent.[21,22]

Information disorder has impacted nearly all facets of society. Globally, people of all ages are spending a significant amount of time consuming and sharing content online. In this way, the magnitude of today's crisis of information disorder is directly correlated to the proliferation of cell phones, increasingly widespread internet access, and availability of social media. The ubiquity of mobile phone access has led to a surge in internet activity throughout the developing world, often through zero-rating that provides internet access without financial cost and offers a limited range of options in terms of which websites users can access.[23]

Figure 2: Groundviews' series on media literacy



Source: https://groundviews.org/2018/05/12/infographic-10-types-of-mis-and-disinformation/

Notably, in Myanmar, Facebook has come under fire for its role in spreading disinformation exacerbating hate speech and contributing to widespread unrest and violence. In fact, the Office of the United Nations High Commissioner for Human Rights called out Facebook for its role in propagating hate speech and inciting the violence that led to the 2017 genocide of the Rohingya people, which resulted in the deaths of more than 24,000 Rohingya Muslims by Myanmar's state forces. The hate speech leading to the genocide spread quickly via online channels and was not shut down due to a lack of Burmese-speaking content moderators on Facebook. According to Reuters, Facebook only had two Burmese-speaking content moderators in 2015, despite repeated warnings that the online media platform was contributing to violence against the Rohingya people. In response, the United Nations set up an Independent Investigative Mechanism for Myanmar in 2018 to collect evidence for use in future prosecutions. As reported by *Malay Mail*, "UN investigators said Facebook had played a key role in spreading hate speech that fueled the violence. Facebook says it is working to stop hate speech and has deleted accounts linked to the military including senior army officials but preserved data."[24]

Key to the current debate about the disinformation problem is the role that social media plays as a vector for disinformation. The rise of social media use and online content creation

information has resulted in a media-saturated world, one that requires a fairly high level of critical thinking from users and consumers—i.e., digital literacy and media literacy.

Although social media was designed to connect us, societies were not prepared for its abuse. The prevalence of online and social media has opened the door for bad actors to use coordinated campaigns to promote and manipulate distorted information for their own ends. The proliferation of coordinated inauthentic activity online greatly threatens the free exchange of ideas and information that democracy is built on while simultaneously challenging societies that are still in the process of developing democratic governance.

Both wealthy and developing countries have struggled to adapt to the large amounts and variety of misinformation and disinformation circling on the internet. However, in developing countries, the results can be both life-threatening, as well as detrimental to democratic governance. In extreme cases, misinformation and disinformation has led to violence against ethnic minorities and impacted the outcome of elections.

Annex 3: Emerging Solutions, provides links to some of the key research centers working on information disorder that regularly put out guides, toolkits, newsletters, and webinars. This research focuses on helping differentiate between the real and the fake online. These are useful for those working to debunk false news and promote factual, truthful information.

Civil society can play an active role in countering the type of malinformation that comes in the form of hate speech. In 2014, the Flower Speech campaign (also known as the Panzagar campaign) was launched to counter hate speech in Myanmar in response to the rise in in anti-Muslim violence. The Flower Speech campaign was founded by Nay Phone Latt, executive director of Myanmar ICT for Development Organization, who was himself sentenced to more than 20 years in prison in 2008 for blogging about the 2007 Saffron Revolution. In the flower campaign, hate speech is countered by efforts to promote responsible use of social media and raise awareness of the implications of online behavior. Through partnerships with local graphic designers and Facebook to create a set of positive 'digital stickers' that users can share on the social media platform, the movement has led to some users posting photos of them holding flowers to further spread the message of peace.

Source: Frontier Media, https://www.frontiermyanmar.net/en/profile-the-flower-speech-movement/

A. WHY PEOPLE USE DISINFORMATION

According to research published by *Psychological Science in the Public Interest*, some reasons that use of misinformation is so rampant include:

- *Rumor:* Societies have struggled with the misinformation-spreading effects of rumors for centuries, if not millennia—what is perhaps less obvious is that even works of fiction can give rise to lasting misconceptions of the facts.
- *Politics:* Governments and politicians can be powerful sources of misinformation, whether inadvertently or by design.
- *Vested Interests:* Corporate interests have a long and well-documented history of

seeking to influence public debate by promulgating incorrect information. At least on some recent occasions, such systematic campaigns have also been directed against corporate interests, by nongovernmental interest groups.

- *Media Fragmentation:* Though the media are, by definition, seeking to inform the public, it is notable that they are particularly prone to spreading misinformation for systemic reasons that are worthy of analysis and exposure. The internet and the growing use of social networks have fostered the quick and wide dissemination of misinformation. The

fractionation of the information landscape by new media is an important contributor to misinformation's particular resilience to correction.[25]

Additional reasons information disorder is on the rise include:

- The erosion of trust in institutions, especially government and media institutions.[26] Environments where trust is low are ripe for the spread of disinformation.
- Misuse of technology, through bots and cyborgs spreading disinformation.[27]

In the political sphere, the ability to win elections is now correlated with a political actors' capacity to manage social media platform messaging. For example, Ukraine's staggering landslide election of both President Volodymyr Zelenskyy and his Servant of the People Party in 2019—sweeping away 80 percent of all Members of Parliament (MP)—was based on almost no concrete policy formulations. Zelenskyy built a formidable campaign machine based on social media and a fictional characterization he embodied (as Ukraine's president in a comedic television series) that his opponents could not match.

This use of disinformation has played a central role in the deterioration and backsliding of democracies around the world.[28] Governments and others who propagate information disorder have created social fissures, contributing to a breakdown of public trust in government institutions and media. This effectively destabilizes and fractures the civic institutions that once upheld/demanded transparency and accountability in political discourse. As the 2020 Edelman Trust Barometer Report finds, 76 percent of a global data sample agreed with the statement, "I worry about false information or fake news being used as a weapon."[29]

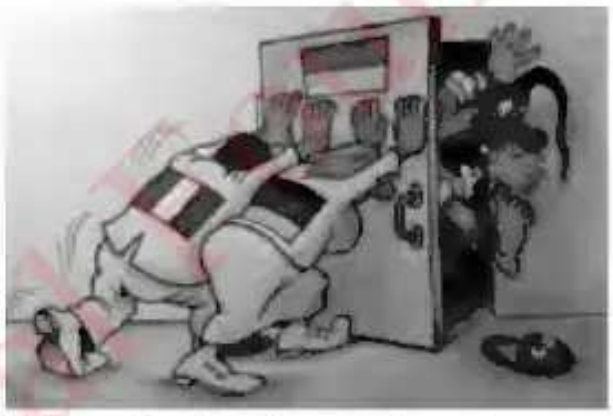


Figure 3: Russian produced meme to persuade Ukrainians against Euro-integration

Source: EuroMaidan Press -- http://euromaidanpress.com/2017/12/15/a-guide-to-russian-propaganda-part-4-russian-propaganda-operates-by-law-of-war/

B. How to identify disinformation

Figure 4, below, was developed by ProQuest to assist in the steps for identifying disinformation.[30] It is very useful to consider when confronted with a news article, social media post, or email that contains claims that seem dubious. While being adept at spotting the different types of false content takes practice, it does become easier. (See Annex 3, Emerging Solutions, for more resources on learning how to identify disinformation.)

C. Context and social factors of disinformation

In assessing a society's risk for problems associated with information disorder, it is vital to consider several contextual and social factors, such as issues of age, gender, and other

marginalized groups; access to technology and the internet; levels of political polarization in society; and legal norms related to content regulation that may impact how issues like hate speech are dealt with.

Rise of digital technology

Digital technology has supplanted many of the forms in which people traditionally obtained information. With the rise of the number of platforms and use of digital technology, consumers now remain more exposed to disinformation. Trends that contribute to this challenge are:

Digital technology has become the norm among people across the globe:

- Global internet penetration is currently at 59 percent.[31]
- More than five billion people have mobile devices, and over half of these connections are smartphones: a median of 76 percent across 18 advanced economies surveyed have smartphones, compared with a median of only 45 percent in emerging economies.[32]

Figure 4: Step-by-step guide to combatting disinformation

1. Do a Visual Assessment

Assess the overall design. Fake news sites often look amateurish, have lots of annoying ads, and use altered or stolen images.
Overall, does the news article and website seem high quality?

2. Identify the News Outlet

The Wall Street Journal and CNN are examples of news outlets. If you haven't heard of the news outlet, search online for more information.
Is the news outlet well known, well respected, and trustworthy?

3. Check the Web Domain

Many fake news URLs look odd or end with ".com.co" or ".lo" (e.g. abcnews.com.co) to mimic legitimate news sites.
Does the URL seem legitimate?

4. Check the "About Us" Section

Trustworthy news outlets usually include detailed background information, policy statements, and email contacts in the "About/About Us" section.
Does the site provide detailed background information and contacts?

5. Identify the Author

Fake news articles often don't include author names. If included, search the author's name online to see if he or she is well known and respected.
Does the article have a trusted author?

6. Identify the Central Message

Read the article carefully. Fake news articles often push one viewpoint, have an angry tone, or make outrageous claims.
Does the article seem fair, balanced, and reasonable?

7. Assess Spelling, Grammar, and Punctuation

If the article has misspelled words, words in ALL CAPS, poor grammar, or lots of "!!!", it's probably unreliable.
Does the article have proper spelling, grammar, and punctuation?

8. Analyze Sources and Quotes

Consider the article's sources and who's quoted. Fake news articles often cite anonymous sources, unreliable sources, or no sources at all.
Does the article include and identify reliable sources?

9. Find Other Articles

Search the internet for more articles on the same topic. If you can't find any, chances are the story is fake.
Are there multiple articles by other news outlets on this topic?

10. Turn to Fact Checkers

FactCheck.org, Snopes.com, PolitiFact.com are widely trusted fact checking websites.
Do the fact checkers say the news story is true?

Source: https://blogs.proquest.com/wp-

In many societies, the amount of time spent online has skyrocketed:

content/uploads/2017/01/Fake-News1.pdf

- In 2019, the average time spent on the internet was 10 hours in the Philippines, 9.3 hours in Brazil, 8.25 hours in South Africa, 7.5 hours in the U.A.E., 6.3 hours in the United States, and 5.5 hours in China.[33] [34]

The growing dominance of social media as a source of information is happening all over the world:

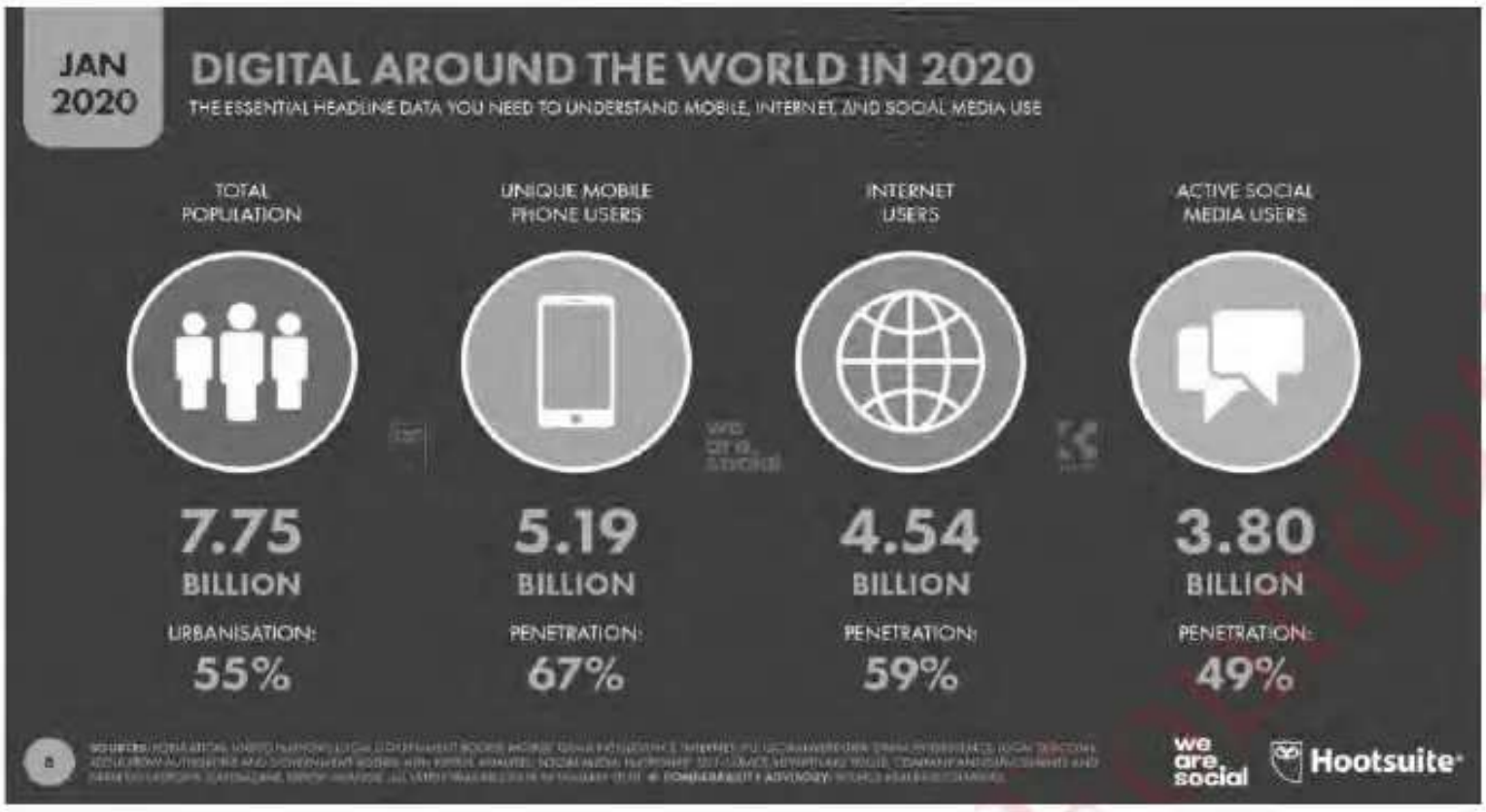
- On Twitter, a total average of 500 million tweets are posted daily; on Facebook there were 1.63 billion daily active users in September 2019.[35]
- India alone is home to 290 million Facebook users. To put this into perspective, if India's Facebook users were a country, its population would rank fourth in the world.[36]
- More than 70 percent of internet users in Kenya, South Africa, Bulgaria, Chile, Greece, and Argentina get their news from social media.[37]

The pervasiveness of Facebook's Free Basics Internet.org—which provides a pared-down cell phone internet experience providing access to mainly social media—has affected internet usage. Social media is becoming synonymous with the internet:

- In many countries, such as Sri Lanka and the Philippines, "opening the internet" on a digital device means opening Facebook. In 2014, Quartz found that 65 percent of Nigerians and 61 percent of Indonesians surveyed agreed with the statement: "Facebook is the internet."[38]
- The Philippines, for example, is a good illustration of how Facebook has penetrated into the social structure: "Free Basics was launched in the Philippines in 2013. By 2018, almost two-thirds of the country's 110 million people were using Facebook, according to *Buzzfeed*. In the Philippines, the word 'Facebook' is interchangeable with 'internet,'" writes Maria Farrell.[39]

Though the platforms may change, the problems that social media brings with it remain the same. WhatsApp, now owned by Facebook, is also making significant headway globally. WhatsApp remains a widely used platform outside of the United States for information sharing. However, WhatsApp's encrypted, and non-public nature makes it difficult to research and analyze.

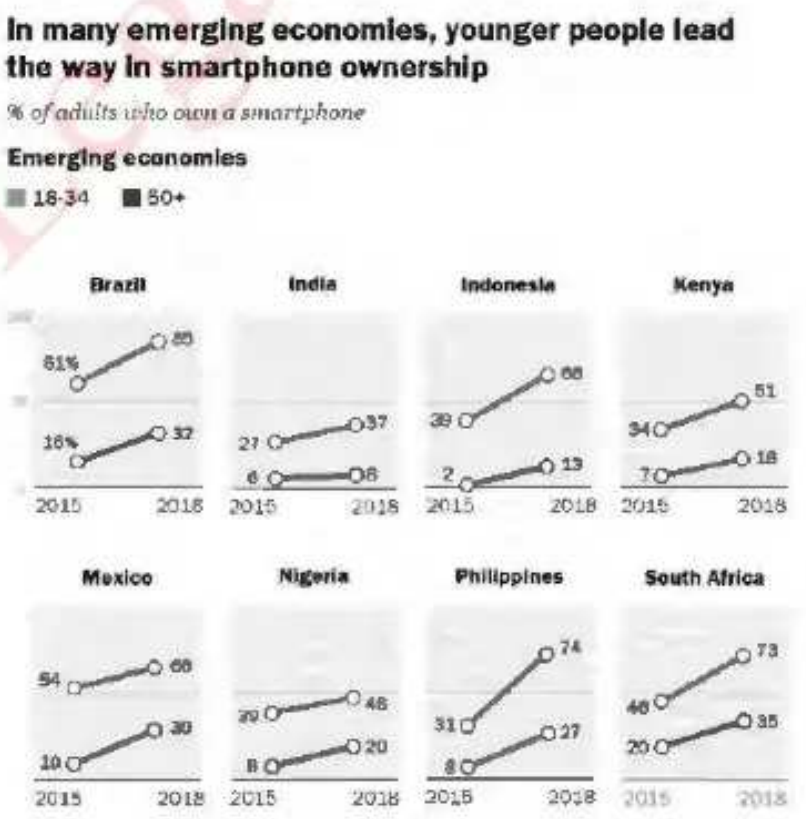
- According to a survey conducted by Reuters in 2017, WhatsApp has become one of the leading news sources in Brazil, Malaysia, and Spain, nearing 50 percent of the population who say they use it for their main news source on a regular basis.[40]
- According to Digital Information World,[41] WhatsApp has 1.5 billion users from 180 countries, which makes it the most popular instant messaging app worldwide. (Facebook Messenger is in second place with 1.3 billion users.)
- WhatsApp has one billion daily active users. The biggest market for WhatsApp in India with over 200 million users; Brazil has 120 million users.



Cross-Cutting Issues

USAID has several cross-cutting

USAID has several cross-cutting issues, such as gender and youth, that are important to link with how people choose to counter and prevent disinformation in their society.



AGE

One of the key questions that needs to be considered in terms of countering and preventing disinformation is whether there are generational differences in terms of news consumption, susceptibility to false news and information, and access to and ability to use technology. For instance, around the world, younger users with a higher education level and higher income are more likely to have smartphone access than their elders.[42] However, statistics show that older people have rapidly caught up and, when they cannot, they often rely on younger users to access the information they need. Moreover, some studies have shown that people aged 65 and older are almost four times more likely to share false news on social media than younger people and that in some instances they are more responsible for the spread of disinformation.[43]Social science research is increasingly interested in the question of whether the consumption of false news is a matter of generational differences. One study found that age plays a key role and has a strong effect on the dissemination of false news. According to

the study (Guess, et al.), "on average, users over 65 shared nearly seven times as many articles from fake news domains as the youngest age group."[44]

In a U.S. study conducted by College Reaction, 69 percent of Gen Z (born mid-1997 to early 2012) students claimed it is somewhat or very easy to discern between true and false information online.[45] However, a majority of middle schoolers in the same generation could not determine the difference between an advertisement and a news story, while 30 percent of the surveyed students found a fake news story to be more credible than a real story.[46] The U.S. experience, however, differs drastically from youth in Finland where children in primary and secondary school are taught about media literacy as a core part of their education, and beneficiaries from Finland's whole of society approach to the disinformation problem. Finland's government launched an "anti-fake news" initiative in 2014 "aimed at teaching residents, students, journalists and politicians how to counter false information designed to sow division."[47]

Some studies have shown elders, who may have less facility with digital technology, to be more immune to disinformation because they rely on other forms (books, education, experience) to assess the validity of claims.[48] Still this bias, coupled with a decrease in memory, may also hinder their ability to discern disinformation.[49]

GENDER

In the Global South, women and men often experience digital technology very differently; however, they use it almost equally in both advanced and emerging economies.[50] Where resources are thin, families often do not have the time or money for women to have access to the internet. Moreover, it is often true that women often have less access to the internet because of local gender norms.

Furthermore, in the spread of disinformation, gender has often been exploited by autocrats and those in power to discredit journalists and eliminate government critics.[51] Female

journalists across the Middle East have been repeatedly targeted with doctored photos of them in sexually compromising positions, claims that they achieved their jobs by being sexually promiscuous, and online defamation campaigns that utilize misogynistic language to discredit them and their news coverage.[52] For more information on this, especially on how women politicians are disproportionately affected by false news, see the Council on Foreign Relations report on Gendered Disinformation, Fake News, and Women in Politics.[53]

In India, journalist Rana Ayyub was slandered by fake images and tweets insinuating that she was a porn actress.[54] The coordinated attacks occurred after Ayyub began fighting for the justice of an 8-year-old girl who was raped over several days and then murdered. The

doctored video of Ayyub was shared over 40,000 times, including a share by the ruling nationalist Bharatiya Janata Party's (BJP) fan page. It ultimately sent her to the hospital for heart palpitations and anxiety.[55]

Stories like this are examples of an overall climate that contributes to the discrediting of female journalists, violence against them, and the danger of covering women's issues in general. A *Reporters Without Borders* report on violence toward journalists covering women's issues found that of all forms of violence and retaliation against these journalists about 40 percent is cyber harassment specifically.[56] It is also worth noting that online violence often accompanies physical violence.[57]

Hate Speech and Dangerous Speech

Hate speech and dangerous speech are considered a potentially life-threatening aspect of information disorder.

The definition of **hate speech** is often contested, particularly because it is such a charged topic, and legal and social organizations offer alternative definitions for the same act. This topic is particularly controversial because there tends to be fine lines drawn in democratic societies regarding what is considered acceptable free speech and what is not. Some definitions consider hate speech a verbal attack made on a group based on a shared identity while others agree that an attack on an individual can be considered hate speech.[58] Likewise, some definitions insist that specific identity markers must be included in hate speech (such as membership in an ethnic or social grouping) while others address any attack on identity.[59]

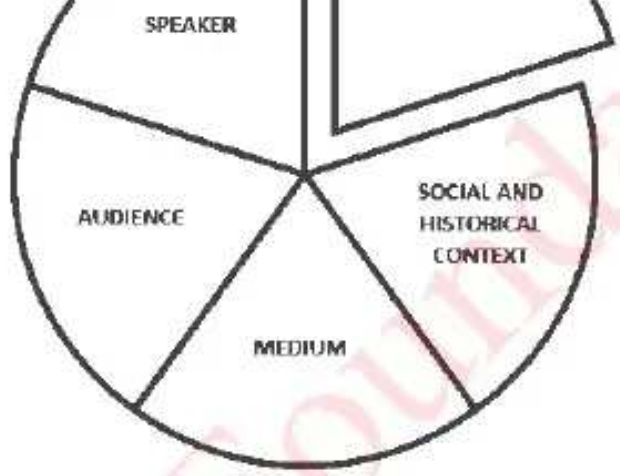
PeaceTech Lab defines hate speech as a deliberate attack or expression that vilifies, humiliates, and discriminates against others based on their ethnic, national origin, religious, racial, gender, sexual orientation, age, disability, or other shared identity. This can lead to a larger societal impact influencing acts of violence.[60] Hate speech is rooted in larger social

grievances that are potentially incendiary and often lead to serious violence and injustice and new research indicates hate incidents online and offline peaking in tandem.[61]

Another term, ***dangerous speech***, is any form of expression (speech, text or

Figure 5: The dangerous speech five-part framework

imaged) that can increase the risk that its audience will condone or participate in violence against members of another group. The Dangerous Speech Project has put forward a framework that considers the speaker, the audience, and the medium, as well as the message and context in determining risk—which is an interesting comparison with the distinction between misinformation and disinformation (i.e., intent versus other factors). Disinformation is also explicitly a category of dangerous speech, where it has the capacity to lead to violence and other types of harm.[62]



Source: https://dangerousspeech.org/guide/

Understanding how to recognize hate speech and dangerous speech is particularly important to combatting their spread online through platforms like Facebook and Twitter, where hate speech and disinformation are often tightly linked.

It is worth noting the concerns about hate speech in the context of conflict (pre or post). A noted example is the Radio des Milles Collines in Rwanda, which called for direct attacks against the Tutsi minority and for which the journalists of that radio were called before the Hague and charged with crimes against humanity. Hate speech, particularly in fragile states marked by conflict, can lead to violence and other harm so it is essential to understand the challenges for labeling hate speech as such.

D. Foreign-Supported Information Disorder

The geopolitical dimensions of information disorder are another key contextual factor to understand. Disinformation is a long-practiced means of statecraft. One of the common forms of disinformation comes from governments and political actors that are working to gain influence. Both the Government of Russia and the People's Republic of China have used disinformation tactics to misinform, change narratives, and accumulate further support for their foreign policies. However, the respective governments approach information warfare in very different ways. One key difference is that Russia focuses primarily on information manipulation while China employs censorship and other forms of information control to suppress other viewpoints and control the narrative.[63]

> "We've seen quite a significant uptick in misinformation generated by foreign state actors, particularly from Russia and China," according to Dr. Phil Howard of the Oxford Internet Institute. "In fact, 92 percent of the misinformation from state-backed agencies around the world originates from Russia and China."
>
> —From CBC News article "COVID-19 disinformation being spread by Russia, China, say experts," msn.com/en-ca/news/world/covid-19-disinformation-being-spread-by-russia-china-say-experts/ar-BB14B35i

While the Kremlin's disinformation campaigns may appear focused on creating chaos for the United States and other Western countries, the reality is that it hopes to utilize disinformation

to weaken perceived adversaries in order to achieve strategic goals, including restoring Russia to great power status, preserving its sphere of influence, protecting the Putin regime, and enhancing its military effectiveness.[64]

The Kremlin's disinformation methods

Kremlin-supported propaganda, disinformation, and information manipulation primarily relies on its advanced network to spread easy-to-understand messages that exemplify a clear narrative of the United States as the aggressor and Russia as the only country brave enough to stand up to U.S. hegemony.[65] Russia has been utilizing strategic disinformation tactics since the 1950s to try to influence the perceptions of people worldwide. The Russian government has for years included disinformation and misinformation as part of "active measures," or covert influence operations.[66] In fact, the word "disinformation" is derived from the Russian term *dezinformatsiya* (дезинформация).[67] Highly coordinated disinformation campaigns are meant to influence the perceptions and actions of others and make it highly difficult to discern between the real and the phony. The goal of these campaigns is to weaken perceived Russian adversaries, often by fracturing the bonds of societies in order to try to weaken or cripple alliances in the West, with the goal of ultimately making it possible for Russia to outcompete the United States and Europe.[68] (Note: It is not just Russia versus the United States; Russia also uses disinformation against Europe and others, and they used it extensively in Ukraine to try to influence events.)

Factory of Lies: The Russian Playbook is an NBC explainer on how Cold War tactics have continued to be used by Russia as a way of subverting the media.

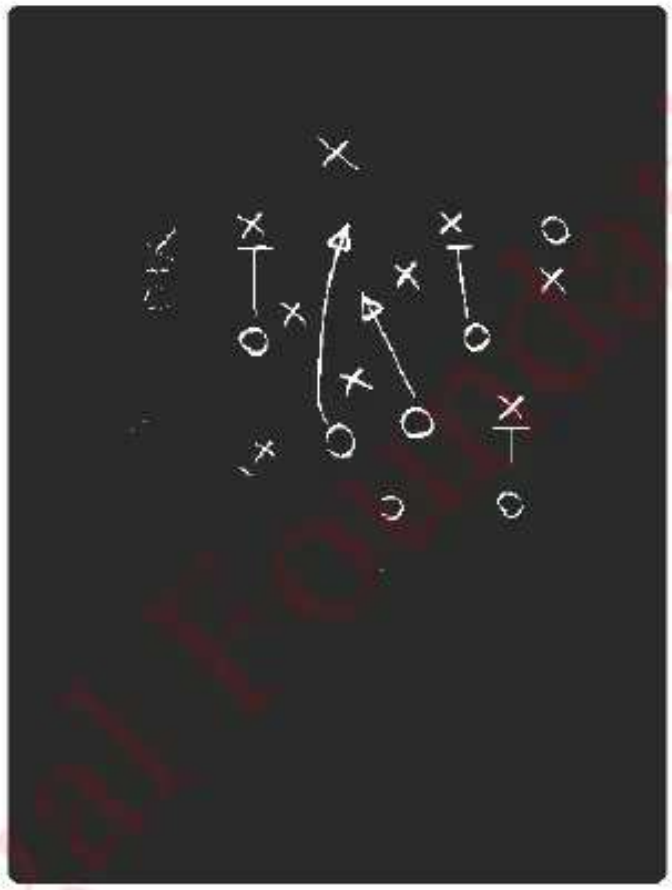
Available at: https://youtu.be/hZrIZU-uZqU

Since 2015, the Kremlin has begun expanding its media influence by creating media cooperation agreements with over 50 local media organizations around the world.[69] The messages shared through these networks play on strong emotions and tend to do very well on social media, where users have been shown to interact with content that is emotional in nature.[70] Thus, in countries where both the United States and Russia have been working to develop influence, the Kremlin tends to put forth narratives that are easy to understand, play to the emotions, and disingenuously offer a clear good guy–bad guy paradigm.[71] Comparatively, the United States has often struggled to offer effective fact-based alternatives to such narratives. This is particularly relevant in countries where USAID works to promote democratic governance.

A key example that merits consideration is the massive disinformation campaign lodged against the White Helmets, known as the Syria Civil Defense, during the Syrian civil war.[72] Russia in particular used propaganda and other disinformation tactics to sow seeds of doubt about the work of the White Helmets and to undermine their humanitarian mission. As *The Guardian* reported, "The aim was to flood the media ecosystem with falsehoods that would erode trust among supporters, especially donor governments."[73] For another example—one within Latin America—both the governments of the United States and Russia are working to influence the geopolitical situation in Venezuela. In order to counter Russia's influence, the United States must work to provide an effective alternative narrative, while simultaneously being aware of the ways Russia creates coordinated disinformation campaigns and offers a rebuttal to negative reporting. Notably, this challenge of countering Kremlin influence is also felt in Ukraine, Moldova, and Georgia (and other post-Soviet states)—countries that are looking to orient their cultures, politics, and economies away from Russia, but face the challenges of ongoing Russian state efforts to exacerbate polarization and conflict. Strategies for these areas, ripe for exploitation, are often based on the goals of weakening the country or destroying its independent, Western, or democratic resolve.

Rules of the Russian Playbook

Russia's strategy for disinformation is a seven-step process intended to ultimately fracture societies from the inside.



1. Look for cracks and social divisions within the target society.
2. Create a big lie.
3. Wrap the lie around a piece of truth.
4. Conceal your hand (make the story seem like it came from somewhere else).
5. Find a useful idiot (who will take the message and push it to foreign audience).
6. Deny everything.
7. Play the long game, resulting in a major political impact years from now.[74]

The internet and social media have given the Russian government an immediacy and reach that it never had previously to continue spreading disinformation and lies while simultaneously slowly tearing away at the fabric of democracy.

In August 2020, the Department of State's Global Engagement Center published a report discussing how Russia utilizes a variety of tactics and channels to create and amplify disinformation and propaganda.[75] Russia's disinformation and propaganda ecosystem is a collection of official, proxy, and unattributed communication channels and platforms consisting of five main pillars: official government communications, state-funded global messaging, cultivation of proxy sources, weaponization of social media, and cyber-enabled disinformation. The Kremlin's willingness to employ this approach provides it with perceived advantages:

- It allows for the introduction of numerous variations of the same false narratives. This allows for the different pillars of the ecosystem to fine-tune their disinformation narratives to suit different target audiences because there is no need for consistency, as there would be with attributed government communications.
- It provides plausible deniability for Kremlin officials when proxy sites peddle blatant and dangerous disinformation, allowing them to deflect criticism while still introducing pernicious information.
- It creates a media multiplier effect among the different pillars of the ecosystem that boost their reach and resonance.

The four Ds approach

The Government of Russia's strategy for dealing with negative reporting on its actions revolves around four tactics:

1. Dismiss the critic.
2. Distort the facts.
3. Distract from the main issue.
4. Dismay the audience.[76]

This strategy allows the Kremlin to maintain control over the information being spread by virtue of discrediting the individual or organization sharing the information, distorting information to fit their purpose and to support state interests, distracting from the situation at hand where it may be at fault, and launching accusations elsewhere and dismaying the audience by warning that moves that negate state interests will have disastrous consequences for those planning them. These strategies—along with a reliance on Kremlin-controlled media and paid or sympathetic commentators in the West—allow the Government of Russia to spread its messages and influence public perceptions around the world.[77]

Helpful Resource

"Canary in a Digital Coal Mine," a new documentary from the National Democratic Institute, shows how digital activism and collaboration between the government and its citizens in Taiwan has helped withstand the threat of disinformation. The film features Taiwan's civic tech community, led by the organization g0v, which uses open data and public tools to fight for transparency and democracy.

Watch at:
https://www.facebook.com/National.Democratic.Institute/videos/642583006293528/

Chinese Communist Party's Disinformation Methods

The Chinese Communist Party (CCP) deliberately seeks to reshape the current world order to Beijing's advantage. The CCP deploys comprehensive, coordinated, "whole-of-government" influence campaigns to promote and maintain Party narratives domestically and globally.[78]

According to a careful observer, the CCP's propaganda apparatus is a critical component in promoting and maintaining its narrative domestically and globally. Its efforts to use censorship, intimidation, coercion, economic incentives, and propaganda to control the information space are a significant component of its attempts to expand its influence worldwide. This approach to information control actively seeks to downplay concerns regarding China's state abuse and surveillance of Tibetans, Uighurs, and members of other ethnic minorities.[79]

Changing Tactics

There are increasing indications that Beijing is taking a more aggressive approach to information manipulation similar to Moscow. The COVID-19 pandemic has demonstrated that Beijing is increasingly promoting disinformation, pushed out by state media, its officials, and CCP-affiliated social media accounts, bots, and trolls. Beijing also undertook concentrated efforts to push conflicting theories about the pandemic which were intended to sow doubt, deflect blame, and create the idea that the PRC is superior to the United States in responding to international health crises like COVID-19. Observers saw an increasing confluence or convergence of Kremlin, CCP, and Iranian regime false narratives regarding

the pandemic.[80] These three adversaries' state information ecosystems have often converged to spread anti-U.S. disinformation, especially to include spurious claims that the United States caused or exacerbated the COVID-19 pandemic. This convergence appears to be a result of "opportunity," not intentional coordination, but all three actors are more

routinely leveraging the information tools of the others in their campaigns. Also, the Kremlin and the CCP share a common agenda in discrediting democracy and advancing non-democratic governance systems (especially as being more effective in responding to the pandemic).

THE PRC'S "SHARP POWER" METHODS

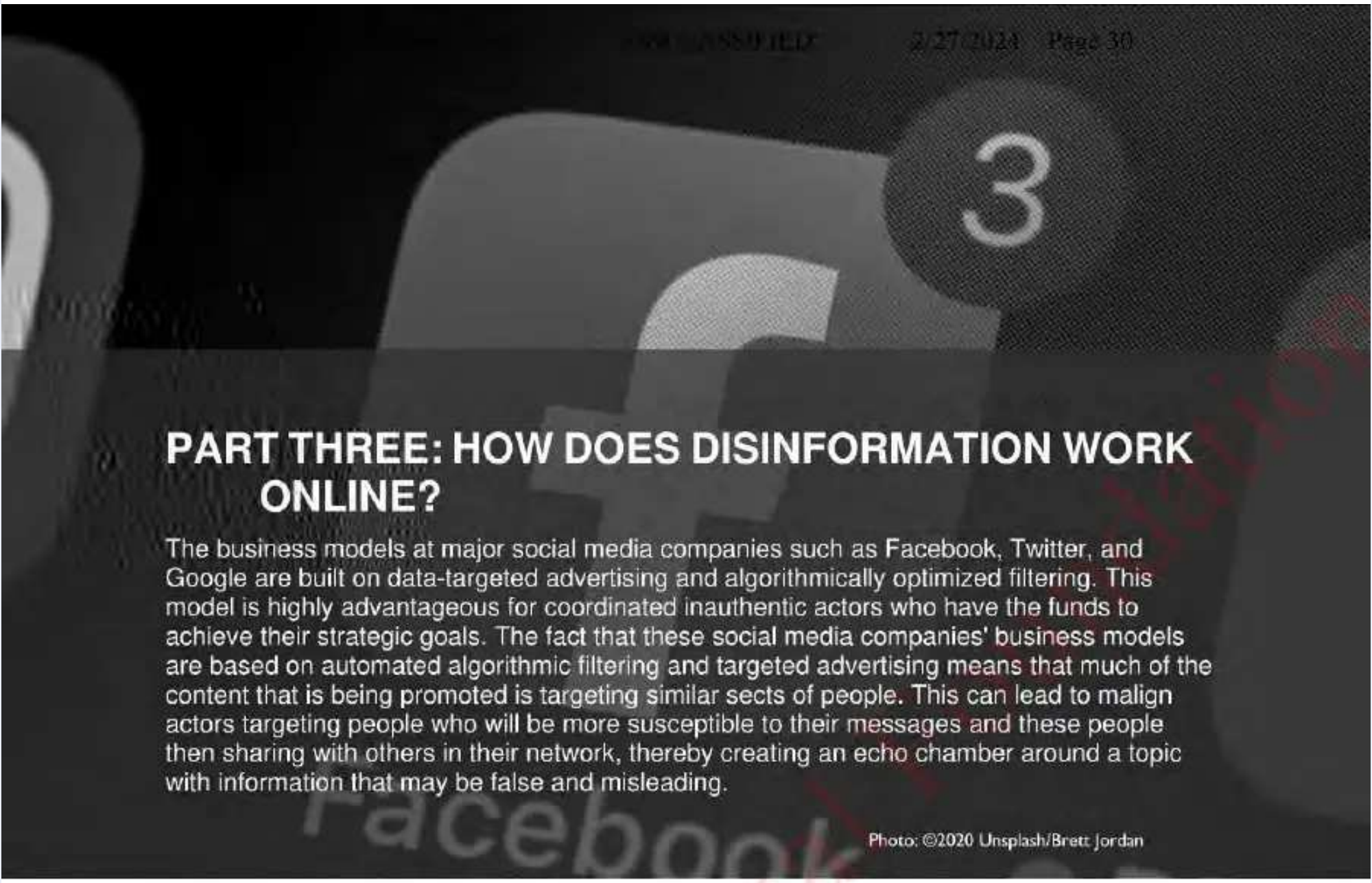
The Sharp Power methods are more about the PRC's control over the narrative about China than disinformation, *per se*. They are a small subset of the CCP's malign influence toolkit. The CCP utilizes all elements of statecraft—such as diplomacy, military, propaganda, and economic levers, in addition to its United Front Work Department (UFWD) —to exert its malign influence abroad. The UFWD, a department under the CCP Central Committee, is responsible for establishing the narratives around strategic issues, especially those concerned with territorial concerns and unification. Major concerns of the UFWD regard ethnic minorities issues. The CCP seeks to "shape and control information flows, bully governments and corporations, infiltrate and corrupt political systems, and disrupt and debase civil institutions," according to the Hoover Institution at Stanford University. These efforts allow the CCP to control domestic and international narratives around Beijing and shape a positive international perception of the PRC to support CCP narratives.[81] These methods are primarily used to control the narrative from what they refer to as the "five poisons:" Taiwanese, Uighurs, Tibetans, Falungong, and pro-democratic activists—the five groups that are most at danger of tarnishing the PRC's global image.[82] Taken collectively, Beijing's "sharp power" strategy seeks to employ "covert, coercive, or corrupting" methods to shape the international world order in areas favorable to CCP interests.

Beijing also has the advantage of platforms like TikTok and WeChat (Chinese-originated applications) that are increasingly used all over the world, which have been used as tools to control and suppress information, especially as it relates to CCP priorities. In particular, the Chinese government blocks information on WeChat and even removes information from private chats.[83] In this way, Beijing controls the narrative about the PRC and maintains a positive image as an alternative to a democratic society. While Beijing has seemingly chosen to focus on the Chinese diaspora, Chinese mainland, Taiwan, and Hong Kong as places to utilize their influencing strategy, the CCP is working to perfect this strategy in order to utilize it in other contexts abroad to influence public opinion.[84]

Across sub-Saharan Africa, the Beijing-based StarTimes television operator provides a popular digital television service, including CCP state propaganda, in the cheapest package available but does not offer alternative international media outlets.[85] StarTimes Vice Chairman Guo Ziqi has stated that their aim is "to enable every African household to afford digital TV, watch good digital TV, and enjoy the digital life." Ultimately, however, this highlights China's strategy of showcasing a positive image toward the world and providing services to developing countries at scale, albeit through the lens of the CCP. Beijing presents itself as a developing country among equals in the developing world and encourages those countries to replicate the CCP's authoritarian governance if they want to pursue economic development without democratization.[86] This view is explicitly intended to offer an alternative to U.S. international leadership and democratic governance. It is also a

central tenet of Beijing's media strategy.

The CCP's investment in media bureaus overseas, content-sharing agreements, and the distribution of content in multiple languages clearly exemplifies Beijing's strategy to influence positive attitudes toward China globally. A 2015 Reuters investigation found that CCP-funded programming was available in 14 different countries.[87] By 2018, a *Guardian* report revealed that the number had grown considerably to 58 stations in 35 countries.[88]



PART THREE: HOW DOES DISINFORMATION WORK ONLINE?

The business models at major social media companies such as Facebook, Twitter, and Google are built on data-targeted advertising and algorithmically optimized filtering. This model is highly advantageous for coordinated inauthentic actors who have the funds to achieve their strategic goals. The fact that these social media companies' business models are based on automated algorithmic filtering and targeted advertising means that much of the content that is being promoted is targeting similar sects of people. This can lead to malign actors targeting people who will be more susceptible to their messages and these people then sharing with others in their network, thereby creating an echo chamber around a topic with information that may be false and misleading.

Photo: ©2020 Unsplash/Brett Jordan

Disinformation thrives best in digital spaces when dissemination agents can construct an effective illusion that changes the behaviors of many authentic users in ways that verify, elevate, and amplify false narratives. Behavior change is sought around things like voting behavior, physical confrontations, conflict, geopolitical orientation, and disruption of democratic deliberation. To better understand the way new technology is used to manipulate social media users and disinform the public, it is imperative to understand some key terms in this field and that this is an evolving space and new types of manipulation are likely to appear.

While there are a multitude of actors working to spread disinformation around the world, research and evidence supports that state actors, and specifically political candidates and national leaders, are increasingly utilizing social media platforms to spread disinformation

about their opponents, manipulate voters, and shape elections. Although in the past, negative campaigning has been utilized in close races and against opponents, the difference now can be seen in the use of artificial intelligence, sophisticated data analytics, and political trolls and bots. Information pollution is increasingly used as a tool to encourage skepticism and distrust and polarize voting constituencies and undermine the democratic process.[89]

The spread of disinformation can occur through manual channels that require manpower, automation, or a combination of both. The Oxford Internet Institute (OII) published a 2019 report that finds "growing evidence of computational propaganda around the world."[90] OII's Propaganda Research Project (COMPROP) investigates the interaction of algorithms, automation, and politics. Their work includes analysis of how tools like social media bots are used to manipulate public opinion by amplifying or repressing political content, disinformation, hate speech, and junk news. COMPROP found evidence of organized social media manipulation campaigns in 70 countries, up from 48 countries in 2018 and 28 countries in 2017.

Manipulation Technique Key Terms and Definitions

Astroturfing: An organized activity that is intended to create a false impression of a widespread, spontaneously arising, grassroots movement in support of or in opposition to something (such as a political policy) but that is initiated and controlled by a concealed group or organization (such as a corporation).

Bots: Social media accounts that are operated entirely by computer programs and are designed to generate posts and/or engage with content on a particular platform.

Clickbait: Something (such as a headline) designed to make readers want to click on a hyperlink especially when the link leads to content of dubious value or interest.[91] This tactic involves creating a misleading or inaccurate post using a provocative headline or image that lures the victim to click and read the content, which is often unrelated or less sensational than the headline itself.

Content Farm: A website or company that creates low-quality content aimed at improving its search engine rankings. Also known as a content mill or factory, its main purpose is to maximize pageviews and revenue generated by advertising on those pages while minimizing the costs and time needed to create the content.[92]

Cyber Troops: Government or political party actors tasked with the use of social media to manipulate public opinion online.[93]

Gaslighting: Technique of deception and psychological manipulation practiced by a deceiver, or "gaslighter," on victims over an extended period. Its effect is to gradually undermine the victims' confidence in their own ability to distinguish truth from falsehood, right from wrong, or reality from appearance, thereby rendering them pathologically dependent on the gaslighter.[94]

Manufactured Amplification: Occurs when the reach or spread of information is boosted through artificial means.[95]

Microtargeting: To direct tailored advertisements, political messages, etc., at (people) based on detailed information about them (such as what they buy, watch, or respond to on a website);

to target (small groups of people) for highly specific advertisements or messages.[95]

Sock Puppets: A sock puppet is an online account that uses a false identity designed specifically to deceive. Sock puppets are used on social platforms to spread or amplify false information to a mass audience.[97]

Trolling: The act of deliberately posting offensive or inflammatory content to an online community with the intent of provoking readers or disrupting conversation. The term "troll" is most often used to refer to any person harassing or insulting others online.[98]

Troll Farm: A group of individuals engaging in trolling or bot-like promotion of narratives in a coordinated fashion.[99]

This research is important to track as it shows the growth in use of computational propaganda (using algorithms, automation, and human curation to purposefully distribute misleading information over social media networks) and social media manipulation by countries, governments, corporations, private actors, civil society, and parties.[100]

The OII research includes a systematic content analysis of news articles on cyber troop activity; a secondary literature review examined public archives and scientific reports, country-specific case studies, and expert consultations. Of the 70 countries where social media manipulation campaigns occurred in 2019, OII broke down the demographics of global disinformation spread as such:

- 87 percent used human accounts.
- 80 percent used bot accounts.
- 11 percent used cyborg (bot + human) accounts.
- 7 percent used hacked or stolen accounts.

The Cambridge Analytica (CA) scandal provides an important case study on the relevance of psychographics and how they are used and manipulated in disinformation campaigns with devastating outcomes for events such as elections.

Of interest to the CA story is how Facebook data could be mined for millions of people and then used to create psychographic profiles. These profiles, in turn, were used for marketing and microtargeting campaigns. According to reports by NBC, "The idea behind the project was that political preferences can be predicted by personal details that people voluntarily provide on their social media accounts. By analyzing the details that users share online, CA could predict individual behavior, which included voter preferences and how to influence that preference."

For an in-depth look at the CA-Facebook scandal, watch the documentary The Great Hack, which shows how illicit use

OII's report offers other key findings that synthesize both demographic and psychographic data about global disinformation spreaders. The findings demonstrate the adaptability of digital disinformation campaigns that can be employed by a variety of players and the prominence of Facebook as the platform of choice when engaging in digital manipulation campaigns.

of personal data can be harvested and used for malign purposes.

Gaining a better understanding of the "who" is imperative to all facets of re-establishing information integrity within an information ecosystem. Knowledge of these profiles, their motives, and favored modes of digital intervention should inform both the information diagnostic process and tactical solutions to suppress the circulation of false news.

A. How algorithms amplify disinformation

An algorithm, as a sequence of instructions telling a computer what to do, is usually built to collect information, recognize patterns, or reach people with particular profiles. Where they can find out how the algorithm works, disinformation agents can craft dissemination strategies to piggyback on a platform's algorithm. In other words, players can game the algorithm to gain access to more digital real estate than would otherwise be possible through simple manpower.

While platform algorithms present users with a litany of navigational and engagement options, they are generally built to elevate and amplify engagement to generate profits. Social media platforms amass profits through paid, targeted advertisement. Through user data analysis, brands can locate specific social media profiles and audiences whose interests,

beliefs, and social behaviors align with their target market audience.

Posts and threads that garner lots of attention and engagement become prime real estate for marketing. Platforms, therefore, are financially incentivized to attune algorithms to amplify posts with the most engagement. Disinformation narratives, troll attacks, gaslighting, and clickbait can generate outrage, opposition, ugly discourse, and/or salacious curiosity that will keep the user's attention and engagement. In this way, computational platform algorithms become accomplices to the dissemination of disinformation because the discourse, which undergirds platform profits, is effectively programmed to elevate sensational content.[101]

Noteworthy in understanding the significance of algorithms: these systems shape the information environment regardless of manipulation from disinformation actors, i.e., YouTube's recommendation algorithm or Facebook's Newsfeed make decisions about what billions of people see by prioritizing types of content they can limit what you see. The key takeaway: bad actors can exploit the algorithm, but without more transparency and accountability we can have bad information outcomes even without bad actors.[102] It is also worth noting the importance of algorithms and the effect they can have on news outlets' business model and bottom line in a digital economy. In the business model of social media platforms' algorithms prioritizing content that gets the most views and interaction, which often prioritizes disinformation and misinformation since they spread more quickly and widely.

B. Coordinated Inauthentic Behavior

Coordinated inauthentic behavior[103] (CIB) is a term coined by Facebook in 2018 to describe the operation of running fake accounts and pages on an online social platform to influence discourse among users. Facebook releases monthly reports that track CIB on its platform and the actions it has taken to address it. As a violation of its community standards, Facebook will remove posts deemed to be taking part in CIB.[104] For the purpose of this

primer, we will be discussing coordinated inauthentic behavior in the context of political actors.

Advanced technologies, such as artificial intelligence, search engine optimization, and bots have allowed for disinformation agents to manipulate information and influence the public sphere using coordinated inauthentic activity. They contribute to a significantly less free internet, where certain voices are amplified above others because of the resources at their disposal and the malicious practices they utilize. Some have called this tactic "censorship by noise," in which artificially amplified narratives and campaigns drown out legitimate dissent.

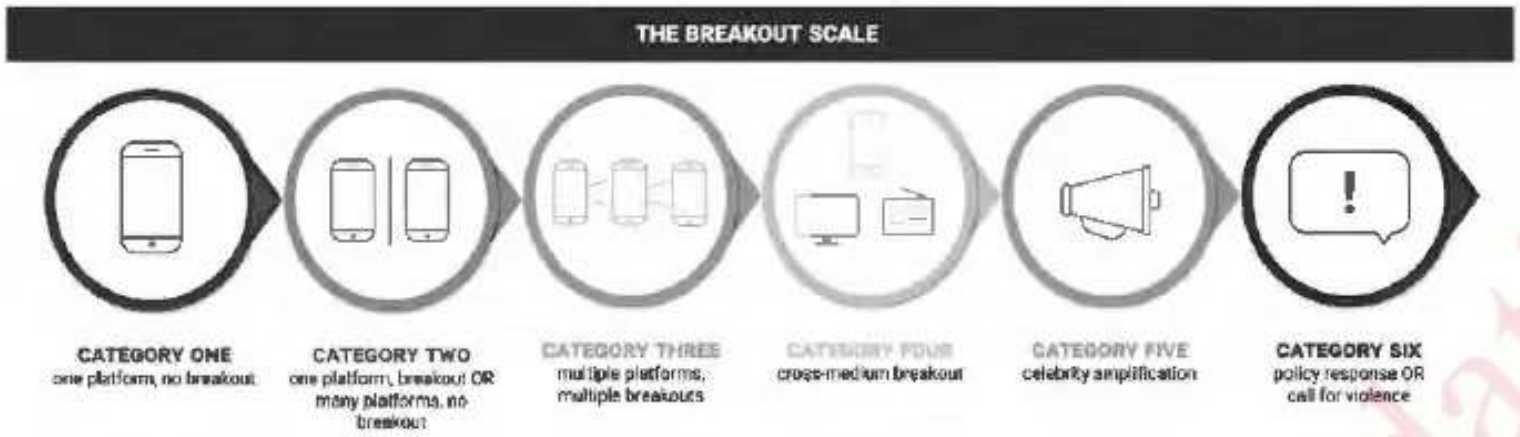
In particular, coordinated inauthentic actors have been able to utilize online systems to their advantage by using algorithms. Algorithms offer a unique opportunity for actors to coordinate messages because many times, the users they hope to influence reside in a filter bubble of content with other users who share similar beliefs and ideals. In turn, these filter bubbles allow content that may otherwise be flagged as inauthentic or false to pass by undetected. As Dipayan Ghosh and Ben Scott argue, targeting and using these algorithms is not an abuse of the platforms as they were designed to market and sell to us; however, the unintended consequences can undermine democracy and spread disinformation.[105]

C. How disinformation spreads across online platforms/applications

Disinformation agents have a variety of advantages when utilizing the internet and social media platforms to share content. Strategies used to mask the originator include the placement, layering, and integration of media messages; this makes it difficult for fact-checking organizations to trace the source of the disinformation. Disinformation agents often obscure their efforts in order to dust off their fingerprints and in doing so work through proxies. While finding the origin of the information is possible, it requires a level of investigative journalism to track it down beyond the capacity of most users and journalists. Additionally, tech companies' advertising models, which are primarily focused on maximizing profit, contribute to the spread. Finally, the organic information ecosystem also makes it easier for information to spread, and the message to be amplified. In sum, the cumulative effective of disinformation has varying degrees of impact, ranging from little to no effect to causing severe harm.

The "breakout scale" (Figure 5) provides useful indicators of the growth of disinformation across online platforms. Each of its six categories demonstrates a growing impact as content travels across multiple platforms (even including offline media and policy debates) and whether it will remain isolated to a few communities or spread through many communities and become a larger or even global phenomenon.[106] The importance of the work by Nimmo and others seeking to map, measure, and understand the disinformation effect on society is that there's an urgent need to put in place preventative measures, including undertaking research studies and engaging in ongoing media monitoring in order to be prepared for the amplification of disinformation that can lead to chaos or worse.

Figure 6: Ben Nimmo's, Breakout Scale: Measuring the impact of influence operations



Source: https://www.brookings.edu/research/the-breakout-scale-measuring-the-impact-of-influence-operations/

DISINFORMATION SPREAD ON SOCIAL MEDIA PLATFORMS[107]

In a few decades, the online media industry has grown from a new frontier with privileged access for tech-savvy groups and individuals to one that supports the most profitable and growing industries. The internet provides a unique advantage for malign actors to spread false content and let it organically amplify as it spreads across platforms. Some conspiracy theories and false claims originate on niche platforms such as conspiracy forums on 4chan or Discord or gaming platforms like Twitch. Platforms such as these enable users to coordinate to grow followers and spread content to large social media sites such as Facebook and Twitter.[108]

In this way, platforms that cater to very niche and small audiences have significant influence. A post shared from one of these original sites can cross to intermediate sites such as Twitter, direct-messaging groups, and Reddit, where they gain traction and are then be amplified further on massive social media platforms such as YouTube or Facebook. Journalists, politicians, and influencers find the content and push it on to even larger audiences. This content now becomes part of the public record and credible news outlets often feel obliged to cover or debunk it, providing it with even more traction. In this way, disinformation goes viral, self-propagating itself from fringe groups to major news outlets. As disinformation takes over, it is an enormous challenge to halt its momentum or inoculate people against it.

The rise of big data analytics, "black box" algorithms (an invisible process through which machines learn about social patterns), and computational propaganda (use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks) have raised concerns from policymakers. The social media business model that includes advertising and data sales promotes controversial posts and information (even though Facebook and YouTube have worked on addressing this lately). These promotion tools have been used in many countries to expand the reach of divisive social media campaigns, to intensify political conflict, and to weaken public trust in the media, democratic institutions, and electoral outcomes. The threats to democracy are further intensified by microtargeting of messages to specific users through sophisticated and proprietary algorithms.

As *Foreign Policy* argued in an article's headline: Disinformation Is Drowning Democracy:

> "In the new age of lies, law, not tech, is the answer. From India to Indonesia to Brazil, democracy is being compromised by online domestic disinformation campaigns from political parties seeking to gain an advantage. Democratic institutions have not been able to keep up and have instead deferred to tech firms, trusting them to referee online behavior. But this is a task far beyond the limited capabilities and narrow

online behavior. But this is a task far beyond the limited capabilities and narrow motivations of companies such as Facebook and Twitter. If the democratic recession is to end, democratic institutions need to create new rules and hold the responsible parties to account.[109] "

DISINFORMATION SPREAD ON CHAT APPLICATIONS

It is also important to recognize the role that chat applications such as Facebook Messenger, WhatsApp, Signal, and Telegram play in the spread of online disinformation. Particularly in countries in Africa and the Middle East and the Global South as a whole, chat applications are an important medium being used to disseminate key political information, activity coordination, and a platform for sharing news and current events. Additionally, many of these chat applications in the Global South host large groups in which entire communities can participate. A 2018 Oxford report on computational propaganda found evidence of social media manipulation campaigns occurring on chat platforms in about a fifth of the countries that were surveyed; many of the countries were from the Global South.[110]

Chat applications are closed - metrics and data about how they work are not accessible. They are therefore hard for researchers to study and difficult for companies to moderate. The applications use end-to-end encryption and offer users security and privacy. Growing in usage globally, these applications have the potential to spread messages to rather large groups. Because of this, some have been limiting the size of groups with whom you can share to address large-scale dissemination of mis/disinformation. Another response to slow the viral spread of mis/disinformation on closed messaging platforms has been to limit the number of times a user can share messages. More on WhatsApp and closed messaging systems is found above in the section on Context and Social Factors of Disinformation (page 13).

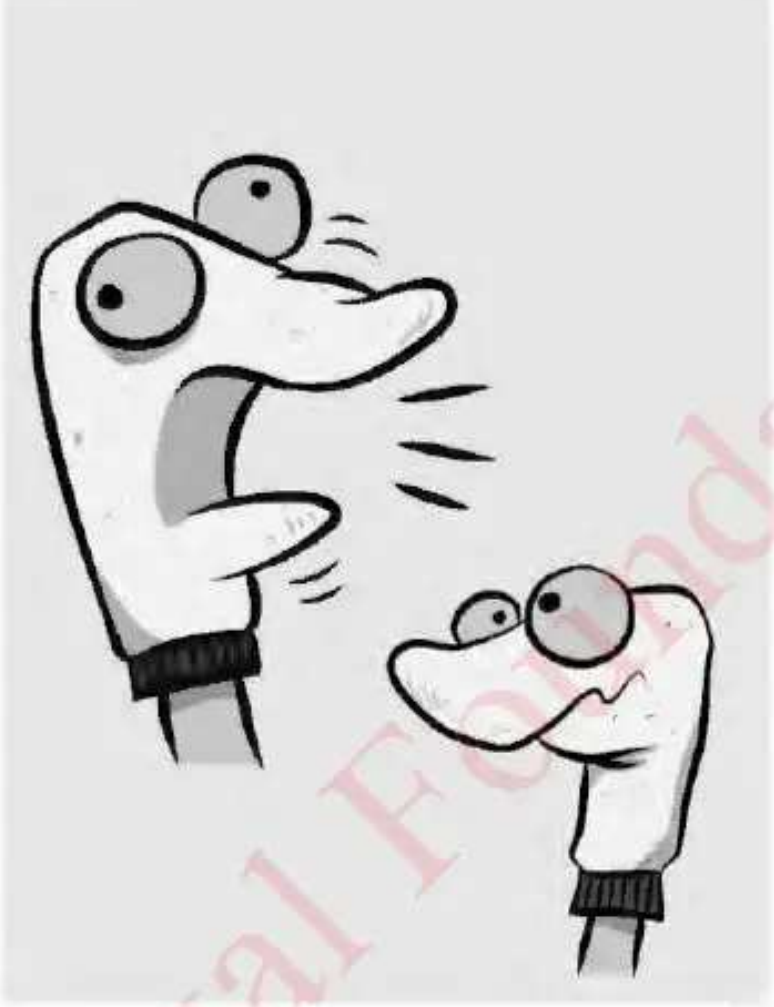
D. TROLL FARMS AND SOCK PUPPETS: CASE STUDY FROM THE PHILIPPINES

A case study that illustrates many of these tactics is the current state of information disorder in the Philippines, widely considered to be patient zero of the disinformation tactics that are now used globally.[111] Because Facebook *is* the internet in the Philippines, content farms, troll farms, and sock puppets have been effective in both the spread of disinformation and the suppression of opposition voices.

Troll Farm: A clear example of a troll farm is Filipino President Rodrigo Duterte's propagated use of trolls to target and defame political opponent and vocal critic Leila de Lima. After pornographic images went viral in 2016, in which de Lima was falsely attributed as the woman pictured, troll farms coordinated by the Duterte election campaign pushed the false narrative within Facebook communities.[112] Trolls used the false content to shame her, attribute her to other scandals, and attack de Lima's character in efforts to delegitimize her as a viable political candidate in the then-upcoming election.[113] Though the video was ultimately exposed as false, de Lima's reputation was stained, and she was arrested on drug charges, which she denies, six months later.[114] Since her detainment, critics have pointed to trolls' spreading of conspiracy theories and misinformation on Facebook to helping lead to her arrest, as well as to distort the public's understanding of the national issue of drugs and further damage the country's democratic processes.[115]

Sock Puppet: The publication *Rappler*

investigated suspicious accounts linked to online Facebook groups and came across "Mutya Bautista," a supposed software analyst at a Filipino broadcast network using a Korean pop star's picture for their profile. Although "Bautista" has only 21 Facebook friends, they are connected to over 160 groups, each with tens of thousands of members. In these groups, the persona and other sock puppet accounts chime into political discussions with real users, repeat false narratives, post politically motivated family anecdotes, and link to false news stories in comments.[116] Sock puppets such as "Mutya Bautista" can be used to support the illusion that false narratives are believed by regular citizens or to drown out those voicing opposition to the falsities in attempts to sway opinion on the perceived power and support behind their position.



IV. PART FOUR: WHAT SOCIAL FACTORS CONTRIBUTE TO DISINFORMATION?

This part explores the social dimension of disinformation, drawing from the latest psychological and sociological research, to consider why we are so susceptible, as humans, to disinformation. It is essential to understand why humans are susceptible to misinformation and disinformation and who is vulnerable. People not only consume disinformation; they call for it themselves, demanding content that they think serves a purpose in their lives, whether it is true or not or in their own interest. How do we understand the factors that drive this consumption of disinformation (the demand side) and make it more likely to be accepted and actionable? And what do we need to know about the social factors promoting the production of disinformation (the supply side)?

Photo: ©2020 Unsplash/Mike Stoll

A. CONSUMPTION OF DISINFORMATION (THE DEMAND SIDE)

A useful—although imprecise—distinction about the drivers of disinformation is that they can be passive or active. Passive drivers are largely subconscious; they require no conscious motivation for individuals to seek out and consume false claims. For example, a person may share information without even reading if it comes from a trusted family member or friend (this is called "familiarity effect;" see Annex 4: Passive & Active Drivers Of Disinformation). On the other hand, active drivers are informed by an individual's thought processes and efforts to understand ideas and reach conclusions through cognitive processes.[118] In this way, a person may believe information that confirms or conforms to a deeply held conviction (confirmation bias).[119]

Some passive and active drivers of disinformation and the reasons disinformation can be psychologically difficult to discern are described below.

1. Passive drivers of disinformation

In general, people are great passive consumers of information that is passed on to them. This tendency is amplified online and can result in many individuals reading and reacting to often emotionally provocative content. Coordinated inauthentic actors rely on this emotional content, to reinforce or encourage people to act. In evaluating what factors lead to accepting information without taking the time to critically engage with it, Woolley and Joseff provide a useful list of cognitive drivers (see Annex 4, Passive and Active Drivers of Disinformation, for more detailed definitions of passive drivers of disinformation) that make it difficult for an individual to discern between truth and falsity and, in turn, make it easier to manipulate them into believing false content.

For example, online content, such as pictures of people in distress, can prime individuals ("priming") by reinforcing existing subconscious biases, thus shaping their perceptions or behaviors. Or by the repetition of a claim ("repeat exposure") many times over, a claim that was clearly false on first reading may be taken as more valid than it is. "Truth biases" such as the tendency to believe printed or online claims on their face may generally serve us well; in the digital sphere, however, they also open us up to even absurd claims.

People tend to overestimate the depth of their knowledge regarding topics they care about. This provides them with an illusion of truth or explanatory depth in the information presented to which they are exposed[120] and may reinforce their beliefs. This "belief perseverance" may be helpful to explain why individuals often remain staunch in their beliefs even after reading contradictory claims. When individuals are asked to think critically about their beliefs and shown information that contradicts them, they often maintain that their beliefs are still correct. Such perseverance of beliefs even when they have been debunked by fact-checking has been thought to create a "backfire effect" in which beliefs are reinforced by the attempt to debunk them. However, some current research suggests that debunking efforts may work well if the facts are relatively unambiguous.[121]

2. Active drivers of disinformation

Active drivers are distinguished by the conscious pursuit of fact claims that serve the purpose of the information consumer. Woolley and Joseff's list illustrates some of the reasons that drive people to seek out false information. (See Annex 4. Passive and Active Drivers for Disinformation, for more detailed definitions of passive drivers of disinformation.)

A common driver is "directionally motivated reasoning" in which people actively consume disinformation to reinforce a specific conclusion they want to reach for political, ideological reasons or in order to reinforce their preexisting opinions (see also "confirmation bias" and "prior attitude effect").[122]

Even if an individual knows that information is false, she or he may be driven to believe it anyway. This often relates to societal pressures that may influence which beliefs individuals publicly adhere to.

An individual might think it is important to believe even dangerous ideas if they are important to a group in which she or he belongs ("bandwagon effect"). Similar to this phenomenon is "consensus bias," in which an individual may believe false claims because of the perception that everybody else believes them and "in-group favoritism" in which specific group identities are at stake.

Often, active consumption of disinformation may require a person to ignore contradictory beliefs or knowledge of facts. This effect, "preference falsification," occurs when individuals suppress their true opinion in favor of societal pressure to support another preference.[123]

The rapid spread of and demand for disinformation may be attributed to laziness or lack of capacity to exercise critical thinking when consuming information. Gordon Pennycook and David Rand, psychologists who have studied and research psychological demand for disinformation, examined the extent to which participants were able to think critically about disinformation in a series of cognitive tests. They showed participants true and false headlines taken from social media from diverse sides of the political spectrum. The results of the study showed that participants who on the first test had shown preference for

utilizing their cognitive reflection did better at discerning the true headlines from the false no matter their political affiliations.[121]

B. PRODUCTION OF DISINFORMATION (THE SUPPLY SIDE)

Anyone with a keyboard can produce disinformation (the supply side) in the form of posts on social media or on the profusion of "news" sites catering to every taste. Disinformation is produced by governments, companies, and individuals to purposefully manipulate public perception and political events.

On social media, fact-checking services are still-developing rules and algorithms are starting to scrutinize this information supply. But even while it is possible to take down or remove the worst posts, for every post taken down multiple mutations of the original disinformation will take their place. Twitter received 27,500 legal requests to remove tweets from July to December 2019 from 98.000 accounts due to suspicious activity.[125] However, this cut-off of supply can have a strangling effect on free expression: as Twitter notes, 193 accounts subject to legal action by governments were of verified journalists and news outlets, and much of the removed information continues to circulate in some form of retweets.

Outrage over "fake news" is overwhelmingly targeted at the creators and distributors of misinformation. That is understandable since they are responsible for releasing half-truths and outright falsehoods into the wild. However, there are indubitably as many drivers or reasons for producing disinformation as there are human interests—albeit political destabilization, ideology, hate, illegal activity, theft, and other criminal activity are age-old motivators. Many observers stress that the supply of disinformation will always be an issue, but we must really focus on the demand side as much as possible.[126]

On the other hand, most suppliers of disinformation are not criminal elements and the ability to understand the motive and the capacity of the supplier to discern the truth are a significant component of digital media literacy. As Pennycook and Rand point out, "Analytic thinking is used to assess the plausibility of headlines, regardless of whether the stories are consistent or inconsistent with one's political ideology."[127] Memes are an important delivery mechanism for producers of disinformation. The disinformation narrative is often couched in the form of memes (an idea that propagates rapidly). The term is now used most frequently to describe captioned photos or GIFs that spread online; the most effective are humorous or critical of society.[128] Memes can be fun to share, but researchers note they also can be dangerous and are a common means through which disinformation, misinformation, and malinformation is spread.[129] They are considered effective because they are

Meme Warfare: Design in the Age of Disinformation

The internet dream has become a nightmare as the information we share is increasingly false and misleading, often with tragic real-world consequences. This video features designer and illustrator Dan Stiles, whose clients range from Arctic Monkeys and Tom Petty to McDonalds and Google, as he examines our role as creatives tasked with creating, collecting, and disseminating information in this radically altered media environment and what we can do to help restore order.

Source: https://www.adobe.com/max/2020/sessions/meme-warfare-design-in-the-age-

catchy, they go viral or have the tendency to spread fast, and they can be shared widely to a large group of followers with relative ease.[130]

of-disinformation-od6303.html

C. EXPOSING DISINFORMATION

When we have evidence of disinformation messages and hate speech and believe they are spreading rapidly, how do we monitor and find out through appropriate research how and how broadly they are circulating? Digital forensics and network analyses as well as traditional media monitoring have emerged as some of the best approaches to track the flow of disinformation. Exposing disinformation is often the first step in countering it.

Social network analysis

Social network analysis is a useful research method both as a diagnostic tool and as a means to develop a strategy for countering and preventing disinformation. Exposing disinformation is really the first thing you need to do. It is the first intervention that must happen and this research can provide an evidence base to inform program design. The analysis of social networks and their role in diffusing new ideas or deepening belief in existing ideas has been advancing rapidly in the last decades. Detailing types of relationships, gaps, and strongly interconnected communities help to understand both the capacity of false information to propagate and the difficulty in correcting it in isolated communities.

Social network "diffusion models" developed originally for epidemiological purposes to track the spread of disease are used regularly to provide graphic maps and algorithms tracking the spread of disinformation. The diffusion model in Figure 6 (below) shows how a network of social media, individuals, and domains enabled the spread of a Kremlin-orchestrated disinformation campaign in Ukrainian elections.[131] Diffusion models have been particularly helpful in understanding the reach and impact of social media platforms such as Facebook and Twitter to trace the sources of information through social media posts and reposts. Social network analysis is a useful research method both as a diagnostic tool and to develop a strategy for countering and preventing disinformation.

Figure 7: Valerij Zaborovskij's diffusion model

A network is a complex system of actors—called "nodes" in graphic representation—each connected by a series of relationships. Information as well as resources can pass between nodes that have established some relationship. A relationship need not be a deep one. In fact, distant relationships, sometimes with people whom we have never even met, can open access to new ways of thinking or to resources that are otherwise out of our reach.[132] In a network graph, the weight or strength of the relationship, the direction of the relationship and the degree or distance of people connected by multiple relationships all can tell us something about how rapidly and effectively a piece of information will pass from one person to many.

The graphs in Figure 7 and 8 shows how a network formed around attempts to influence international perceptions on pro-Indonesia disinformation about the separations movement in West Papua. As Benjamin Strick shared in a blog post, "The campaign, fueled by a network of bot accounts on Twitter, expanded to Instagram, Facebook and YouTube. The content spread in tweets using specific hashtags such as #FreeWestPapua, #WestPapuaGenocide, #WestPapua and #fwpc."[133] The full SNA graph produced by Strick can seem daunting (see below); however, it contains important information. If we zoom in on a section, for example, the graph reveals central nodes—actors who are important to disseminating information widely. Knowing this can improve the focus of countering disinformation programming and help to cut off disinformation flows.

Figure 8: Full SNA on disinformation in West Papua

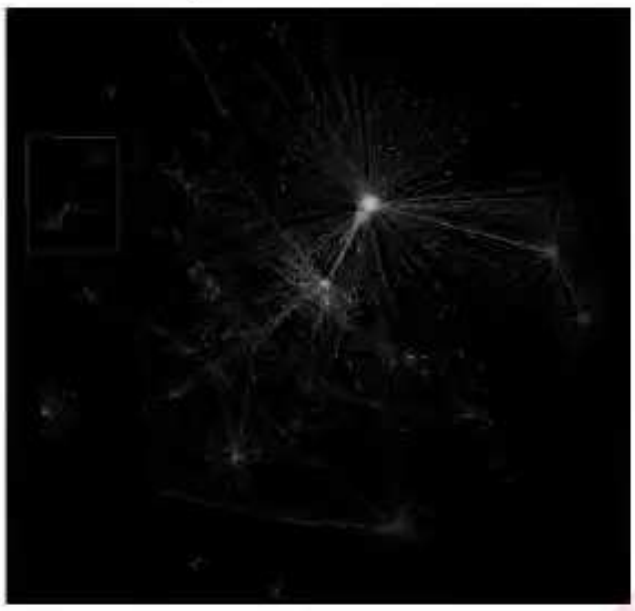


Figure 9: Zoomed in SNA for Papua, showing central nodes

Another social network analysis, conducted by Graphika, linked inauthentic coordinated behavior to influence the 2020 elections in Myanmar to the military who displaced civilian leadership in a coup in January 2021. The analysis tracked numbers of posts, numbers of shared posts, and even the time of day in which posts were shared to reveal the

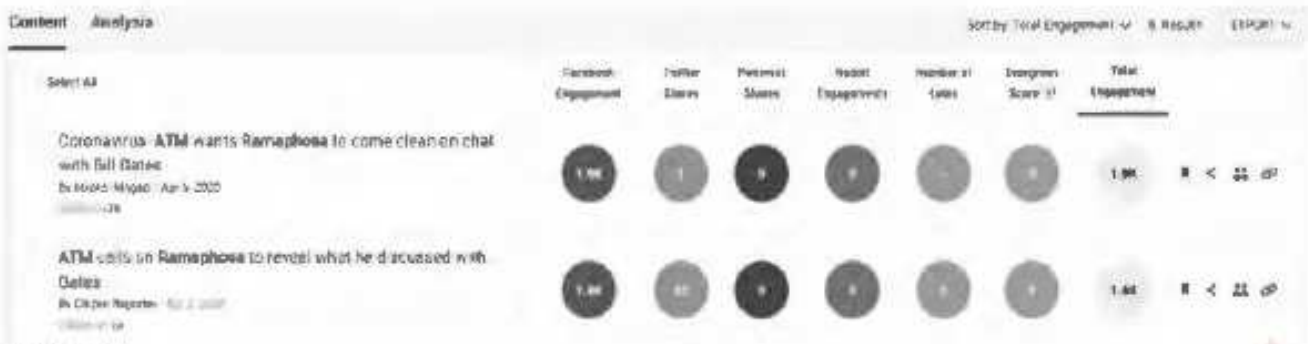
coordination.[134] Facebook removed over 70 accounts. According to Graphika's report that details the social network analysis research:

> As it announced the takedown, Facebook said, "We identified clusters of connected activity that relied on a combination of fake, duplicate and authentic accounts to post content, evade enforcement and removal, manage Pages and drive people to off-platform websites including military-controlled media domains. These accounts often used stock female profile photos and photos of celebrities and social-media influencers. ... We began our investigation after reviewing local public reporting about some elements of this activity. Although the people behind this activity attempted to conceal their identities and coordination, our investigation found links to members of the Myanmar military."[135]

Digital forensics

Digital forensics take a deep look at the data about posts (e.g., number of shared posts). Two of the most-followed fake accounts removed by Facebook this year were focused on Philippine news, according to network analysis firm Graphika. Digital forensics found that these networks originated from individuals in China; Facebook has also shut down another account with links to Philippine military and police.[136]

Figure 9: Forensics on the spread of fake coronavirus information by fringe parties in South Africa



Source: DFRLab

Figure 9 uses digital metrics to quantify the spread of disinformation about coronavirus by South African fringe parties, showing how significant a role Facebook played. Forensics such as these are very helpful in identifying both where and how people are engaged in the digital media ecosystem.[137]

The Atlantic Council's Digital Forensics Research Lab (DFRLab) publishes research along these lines. Looking at how Facebook identified and removed disinformation posts associated with the United Russia party, the research considers the wide scope of data that can be gathered from likes to followers and the types of online activity carried out by suspicious accounts. A look at how these digital networks are connected provides us

Data Analytics for Social Media Monitoring: NDI Guidance on Social Media Monitoring and Analysis Techniques, Tools and Methodologies is a guide to help researchers, election observers, technologists and others understand the best practices, tools, and

with a map that can inform policy that would target the most central players, as represented by the larger circles, representing domains, social media accounts, implementing partner addresses, personas, and Google analytics IDs. Based on this network map, for example, one might expect that a counter-campaign would target and watch for information passing by these means. Facebook, for example, removed 40 specific user accounts, 17 pages, one group, and six Instagram accounts for coordinated inauthentic behavior.[138] Digital forensics (often available as open-source software) provide available data on the backend of internet use—unique users, page clicks, and visits, for example—as critical clues about the spread and origins of mis/disinformation. The approach can help supply data needed for Social Network Analysis (SNA for short) in the digital realm and quantitative information on the reach of websites or other platforms that are needed to inform counter disinformation programming.

methodologies for developing online observation and monitoring for social media networks. It presents an introduction to the relevant concepts when studying these issues, as well as a review of how to build a complete picture of the socio-technical context in a country or region, including the local parties' online presence, social media and internet penetration rates, local media, ethnic and religious divisions, and a host of other factors that manifest in the online space.

Available at:

https://www.ndi.org/sites/default/files/NDI_Social%20Media%20Monitoring%20Guide%20ADJUSTED%20COVER.pdf

The National Democratic Institute has developed robust data collection guidelines to help researchers, election observers, media programs, and others to compile and collect meaningful digital forensics and improve media monitoring efforts.[139]

Media monitoring

While digital forensics focus more on the means and sources of disinformation, media monitoring efforts highlight content issues to understand the impacts and way in which untrue narratives are being constructed. Methods can include social media monitoring, message monitoring, social listening, and more traditional content analysis. All these methods consider the actual words and meanings used to construct disinformation.

Social media monitoring is the process of identifying and determining what is being said about an issue, individual, or group through different social and online channels. It is also used by businesses to protect and enhance the reputation of their brands and products. The method uses bots to crawl the internet and index messages based on a set of keywords and phrases.[140]

Message monitoring analyzes the tropes, narratives or specific messages that a bad actor is putting forward. In this way, it monitors platforms to look at what are the key messages that extremist or conspiracy groups are putting out to see if there are specific messages that they are repeating in talking points. This is a way to understand how individuals are recruited by groups like Al Shabab in Somalia, in order to counter their influence. In long-term programs, such as democracy building or civil society strengthening, it is helpful to look at sources of disinformation and monitor the messages over time.

Social media listening is a means of attaining interpersonal information and social

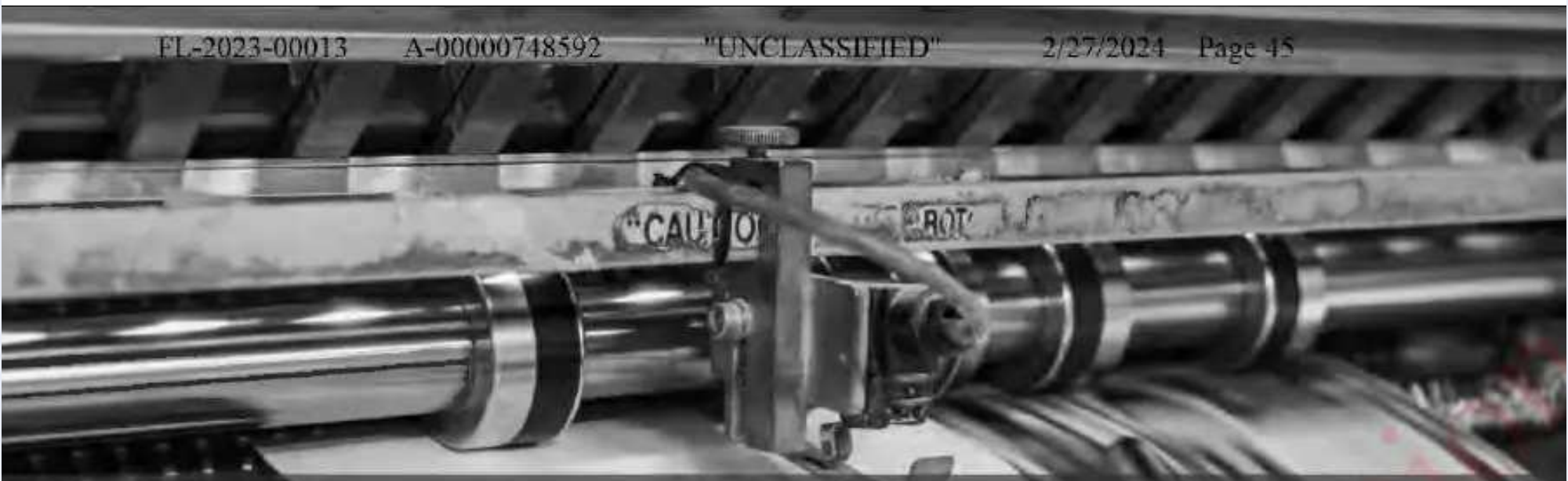
intelligence from social media to understand how relationships are formed and influence the way we listen to and communicate with one another.[141] Fact-checkers can use social listening to develop a more comprehensive understanding of disinformation consumption and groups that might find value in receiving fact-checked articles. Social media listening tools often measure positive, negative, or neutral sentiment. Listening takes perceptions and emotions into account and has been an area of growth among corporations as a way of improving their marketing. Surveys and interviews as well as coding messages for emotional content are ways to listen in on how messages are moving individuals. This can be particularly powerful in campaigns against disinformation because it enables reacting in real time to would-be consumers of disinformation, as well as addressing the issues that make them susceptible to it.

Natural language processing (NLP) is the relationship between computers and human language content. It refers to speech analysis in both audible speech, as well as text of a language. NLP systems capture meaning from an input of words (sentences, paragraphs, pages, etc.). In this way, NLP proponents are working toward a greater capacity of computers to detect fake or fabricated messages that are often couched as satire or hidden within unrelated topics.

For more information, see the Media Monitoring section of Annex 5: Section-by-Section Resources.

Open-Source Intelligence (OSINT) is another strategy utilized for identifying and debunking disinformation. This refers to the multi-method approach of collecting and analyzing free, publicly available information and cross referencing it against other public sources.[142] Publicly available information often includes material from satellite images, social media posts, YouTube videos, and online databases, among other sources.[143] OSINT is noted for its accessibility as a free tool that anyone can use.

For more examples, see the OSINT section of Annex 6: Section-by-Section Resources.



V. PART FIVE: WHAT ARE SOME ANTICIPATED CHALLENGES?

Today's digital communications and media landscape is complex and has given rise to a new set of challenges and considerations for democracy support. Across all USAID programming

countries, this requires a solid understanding of information disorder and robust approaches for countering and preventing it. Note: you will want to make your work as context specific as possible and commission or fund original, diagnostic research to determine the best course of action for your countering and preventing disinformation strategies.

Photo: ©2017 Unsplash/Bank Phrom

Looking forward, combatting disinformation will remain a serious challenge for societies around the world and a danger to democratic governance. As technology continues to outpace the development of solutions, disinformation tactics will adapt and innovate to remain effective. These inevitable cycles of new disinformation techniques and solutions that provide temporary patches are evolving and becoming more sophisticated in the global competition over the control of information.

Future action should concentrate more on critical research and the expansion of knowledge on technology innovation, programming, and information systems. There remains ample opportunity to explore and develop more tech-based tools and approaches. However, to best address disinformation, action and research cannot be left to technology experts alone: the general public and civil society organizations need to gain a basic understanding and grasp of digital and information literacy.

The following sections present a sample of trending disinformation tactics, rising threats, and potential opportunities. As technology continues to innovate and learn from its previous shortcomings, new evolutions of tools and tactics present concern for future information disorder. From the expansion of artificial intelligence capabilities to the exploitation of existing vulnerabilities, anti-disinformation approaches will face new challenges from varying angles.

A. Exploitation of Areas of Declining Media Coverage

The changing media landscape, from the closure of newsrooms and print newspapers to the rise of digital media consumption, has led to the emergence of news deserts: areas in which residents have limited access to news and information outlets.[144] News deserts can include areas where news is unavailable in minority languages, areas in conflict, and areas with a

high degree of state control. The problem of news deserts is a familiar problem in international development: addressing the lack of access to news in the Global South is the focus of the media development sector.

Research from the University of North Carolina found that news deserts tend to occur in communities that are much poorer, less educated, and older.[145] In a period marked by disinformation and misinformation, the rise of

news deserts positions already-vulnerable populations in an even more disadvantageous situation, cutting them off from critical access to education, health, safety, and political information, among other topics. It also has a negative effect on local governance—the management of budgets, elections, and local problem-solving, for example.

Although local newspapers have tried to transition to digital operations, the rise of Big Tech has inhibited their success. With Facebook and Google sharing 80 percent of the digital ad market, smaller organizations are left competing amongst themselves for the remainder of the market, limiting the amount of ad revenue they can generate.[146] Steve Cavendish explains that "print dollars that many news chains have walked away from have been replaced by digital dimes or even digital pennies," leaving them to scale back or close. Ultimately, the rise in news deserts may result in more people turning to social media as their primary sources for news and information. While social media platforms may be widely accessible, they continue to be channels for disinformation, misinformation, and malinformation to spread. These platforms are not a replacement for the institution of a democratic, free media.

Consequently, disinformation actors exploit local news deserts, which has led to a new and growing phenomenon called "pink slime journalism,"[147] a low-cost way of distributing thousands of algorithmically generated news stories, often with political bias. Designed to look like real, local news sites, they are in fact low-cost, automated sites that often push partisan agendas. These pink slime sites capitalize on news deserts left when regional newspapers go broke. While these stories can be factual, they are not based on investigation and may parrot fake claims made in news releases or from opinion leaders. Increasingly pink slime operations are funded by political parties in the United States or by foreign governments (e.g., Iran), highlighting a critical need for transparency.

Pink slime propagators own multiple newsletters and outlets, which enables them to be profitable and makes the niche media essentially a pay-to-play proposition akin to advertising.[148] Because of its diversion from critical analysis, pink slime journalism is an effective megaphone for disinformation.

B. Understanding Alternative Media Spaces

Discussions on disinformation and misinformation often revolve around assumptions of state actors driving the issue. However, problematic information more regularly originates from networks of alternative sites and anonymous individuals who have created their own "alt-media" online spaces.[149] These alternative spaces include message board and digital

distribution platforms (e.g., Redditt, 4chan, or Discord); conspiracy news sites (e.g., RT and 21st Century Wire); and gaming sites. According to Eliot Higgins, founder of Bellingcat, the alternative media ecosystem has become a prominent driver of disinformation, yet not many organizations, journalists, or researchers are engaged in this topic.[150] This present failure to address alternative media systems threatens and undermines other efforts to counter

disinformation.[151]

While information on alternative systems such as conspiracy theories may seem farcical or preposterous to an outsider, to users these spaces enable them to collaborate and validate their own claims and interpretations of the world that differ from "mainstream" sources.[152] With this, individuals contribute their own "research" to the larger discussion, collectively reviewing and validating each other to create a *populist expertise* that justifies, shapes, and supports their alternative beliefs.[153] As these discussions become larger, "mainstream" institutions may pick up on the issue but because they do not understand the platform or alternative media system more generally, they may unknowingly provide wide coverage of misleading information.

C. "NARRATIVE CONTROL" BY STATE ACTORS

Authoritarian and hybrid regimes tend to clamp down on dissenting voices while more democratic regimes struggle to find the best balance in the effort to control the narrative, often justifying these actions as a way of addressing mis/disinformation. The use of legislation and policy, as well as internet shutdowns are discussed below.

1. Illiberal legislation and policy

Trends in legislative action around information disorder issues suggest many governments are under the impression that disinformation can be "legislated away." However, countries who have chosen this route have met many obstacles and criticisms along the way. Outcries around censorship or the broadening of executive powers haunt many governments who attempt to regulate citizen behavior in digital spaces.

In Nigeria, for example, after disinformation campaigns rattled the country's 2019 elections, the Nigerian senate took up a "social media bill" to criminalize the posting of false content on social media if the content is deemed to destabilize public trust in the government, attempt to influence elections, or compromise national security.[154] Critics say the bill will jeopardize digital freedoms of expression while granting the government sweeping, unchecked authority over the country's media environment.

Nigeria's social media bill is almost identical to Singapore's Protection from Online Falsehoods and Manipulation Bill (POFMA). The October 2019 legislation established a nine-member committee to preside over the prohibition of posting politically motivated false statements and the creation of inauthentic online accounts (bots or sock puppet accounts) within digital platforms. The committee can charge individuals or entire

Case study: Nicaragua

Nicaragua passed legislation in October 2020 that will criminalize spreading "fake news" on social media. The Nicaraguan example has sounded the alarm bells for press freedom advocates. Given the political nature and imprecise nature of the term "fake news," such a law would be difficult to apply fairly and is likely to be used to increase political repression. Rights groups believe the law is a specific attempt to silence them.

Source: Lopez, I. (2020, October 27). *Nicaragua passes bill criminalizing what government considers fake news*. National Post, Reuters. https://nationalpost.com/pmn/news-pmn/crime-pmn/nicaragua-passes-bill-criminalizing-what-government-considers-fake-news

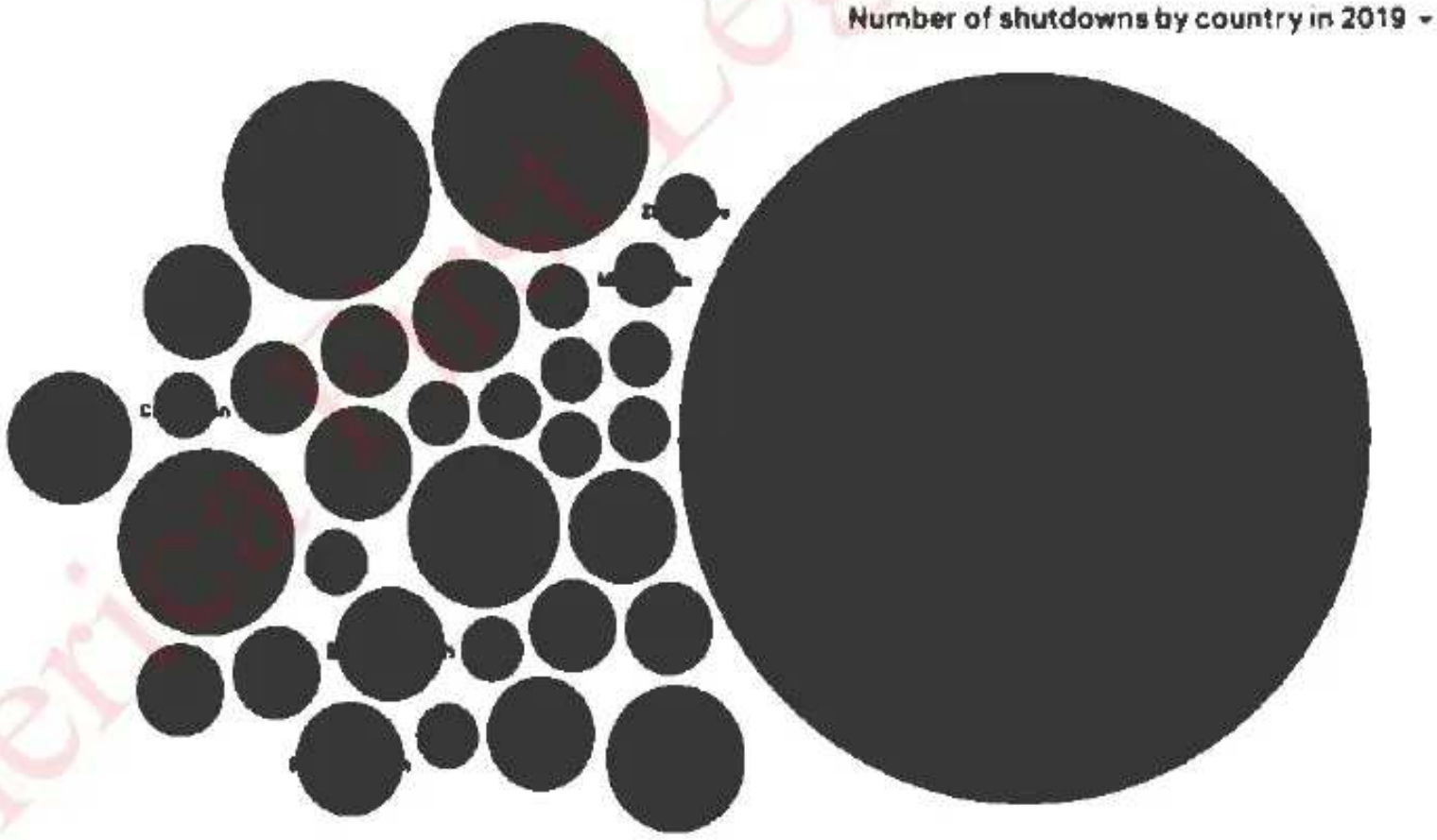
media outlets if content is deemed false or misleading or if implementation of the law serves greater public interest.[155]

Both the Nigerian and Singaporean laws are subjective in nature in that they rely on an interpretation of a user's digital actions and intent. Some critics argue this is impossible to concretely access. Likewise, the gray area of intent can become an easy decoy for governments trying to exercise censorship under the guise of law. The constant swarm of controversy surrounding many such bills often causes them to slow and stall in legislative processes.[156]

Taking disinformation legislation a step further, the Ukrainian government attempted to develop new legislation to criminalize the persistent dissemination of disinformation, create a new ombudsperson to tackle disinformation, and oblige non-governmental media organizations to somehow merge to form a supervisory body that would accredit journalists and determine good-quality from bad-quality media.[157] Ukrainian civil society was united in strongly criticizing this approach, and the government dropped the idea. The weaponization of the disinformation dilemma and consequential chilling effect erodes the integrity of journalism and information within societies.

Some governments believe that internet shutdowns or slowdowns are a solution to the problem; they are not. According to AccessNow, the impact of shutdowns affects journalism and access to information, education, refugees, healthcare, and business, not to mention violates the fundamental right to access to the internet as an essential right in the 21st century.

Figure 11: Number of internet shutdowns in 2019



Source: AccessNow, from #KeepItOn Campaign and research.

AccessNow's *#KeepItOn* campaign reported 216 internet shutdowns worldwide in 2019. These broadband and mobile network disruptions represent 1,706 total blackout days in 33 countries.[158] National and local governments that implement internet or platform blackouts often justify the action as a measure of public safety or national security against the social

harms of fake news. However, international free speech and press freedom advocate organizations denounce blackouts as authoritarian and hazardous during public crises where impediments to current and accurate information is life-threatening. The issue of internet shutdowns is important to monitor because it opens a pandora's box that threatens several areas USAID programming addresses and the use of broadband and mobile networks is often critical to program outreach. Moreover, the economic impact of internet shutdowns is a big deal. They cost $2.4 billion between July 1, 2005 and June 30, 2016, according to the Brookings Institution.[159] And, just a few years leader, the trend worsened. Research firm Top10VPN published a report that analyzed the economic impact of internet shutdowns throughout the world in 2019. Their research traced 18,225 hours of internet shutdowns around the world in 2019 and noted that this carried a total economic loss of $8.05 billion.[160]

The governments of Cameroon and Venezuela have also engaged in internet shutdowns and platforms bans. However, in both countries, the governments have used network access as leverage within existing political conflicts. In Cameroon, English-speaking regions lived without internet access for 240 days in 2017, amid continued civil unrest. Similarly, in Venezuela, President Nicolás Maduro has used internet shutdowns frequently over the last seven years and, in the last year, access to Facebook, SnapChat, Instagram, Google, Twitter, and YouTube vacillated during periods of heavy civilian protest. Both governments have wielded network access to stifle dissent, expand pluralistic ignorance, and silence oppositional voices in digital spaces.

In Southern Asia, where blackouts have become common, the Sri Lanka government disabled access to multiple platforms (Facebook, WhatsApp, YouTube, Instagram, SnapChat, and Viber) after the 2019 Easter bombings, claiming the bans protected citizens from misleading, unverified, or speculative content during the national crisis. The president also pointed at digital platforms as an enabling space for terrorism and hate groups, such as the one responsible for the bombings. The ban had a significant impact on the public in a country where Facebook is often a primary means of both news information and communication with loved ones.

D. DEEPFAKES AND CHEAP FAKES

Deepfakes are videos, images, and audio generated using artificial intelligence to synthetically render realistic depictions of speech and action.[161] Most notably, the technology has been used to manipulate facial expressions and speech, as well as swap the faces of individuals into videos. While altered and manipulated content already circulates online, the development of deepfakes intended to misinform will "significantly contribute to the continued erosion of faith in digital content," according to *Brookings*.[162]

People are more likely to have a visceral reaction to videos, images, and audio rather than text, so deepfakes propel a more rapid spread of altered media.

Cheap fakes are audio-video manipulations created with cheaper, more accessible software (or none at all).[163] They may be subtle, such as slowing down the speed at which a video is played making it appear that the speaker's speech is slurred or altering the background or an otherwise insignificant aspect of a picture.

Many Artificial Intelligence tools, including deepfake-specific technologies, have free and open access, enabling the creation of fake content to expand readily.[164] In the future, cheap fakes are likely to be uploaded by amateurs with satirical or political motives and influence campaigns of foreign origin or with advertising goals. However, despite their aim, cheap fakes' wider proliferation in online spaces will further blur the authenticity of the digital world.[165]

To date, ongoing research and mitigation efforts have concentrated on automated deepfake detection, developing algorithms to recognize modified media. Academic studies have also uncovered indicators of deepfakes, including unnatural blinking patterns, distorted facial features, lighting inconsistencies, and more.[166] However, as deepfake technology continues to rapidly improve, these efforts are likely to be short lived. More funding and research are needed to support the discovery and development of longer-term solutions and tools to detect deepfakes.

E. Artificial Intelligence (AI)-Generated Propaganda

Previously, dubious media outlets, articles from nonexistent authors, troll factories, and comment armies have been deployed from Russia, Poland, Philippines, and elsewhere to saturate online spaces with fake content in order to manipulate public opinion.[167] However, because writing articles, comments, and social media posts can be time consuming, those behind these operations often gave away indicators of their illegitimacy with their rapid activity, including plagiarizing or recycling writing, using stolen profile photos, and using repetitive phrasing in high volumes of text.[168] While these influence campaigns have been unraveled and attributed, advances in AI-generated content could eliminate these tells, leaving them untraceable.

Emerging as the ideal tool for propaganda, AI-generative text resolves the time and effort needed for original content production. Artificial-intelligence research lab OpenAI has

already released a beta version of GPT-3, a long-form text generator that works by taking text input and predicting what should follow.[169] This tool reportedly can produce long-form articles, tweets, poems, and other texts that are difficult to distinguish from a human's writing. Tools like GPT-3 present yet another challenge for internet platforms and users to discern what and whom to trust.[170] If people are left questioning whether content is tied to an actual person or an AI-generator, new divisions over verification of users and moderation on

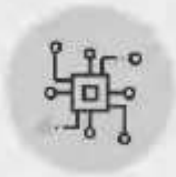
platforms could also emerge.[171] Those desiring more assurances of legitimacy may call for increased account verification or physical existence validations on sites. Meanwhile, those in disagreement with increased moderation could turn to the alternative media systems with minimal restrictions.

VI PART SIX: WHAT ARE SOME EMERGING SOLUTIONS FOR DISINFORMATION?

Emerging solutions for combating disinformation are coming from many different directions across sectors, including approaches to counter violent extremism, prevent hate speech, address national security and foreign interference, improve journalism, engage academia and education organizations, and enhance political communications, among others.

Photo: ©2020 Shutterstock/EriAmmos

Wardle and Derakhshan in their disinformation report for the Council of Europe provide recommendations targeted at technology companies, national governments, media organizations, civil society, education ministries, and funding bodies, with each broken into its own set of suggestions.[172] (See table to find the recommendations.)



What could technology companies do?

- Create an international advisory council.
- Provide researchers with the data related to initiatives aimed at improving public discourse.
- Provide transparent criteria for any algorithmic changes that down-rank content.
- Work collaboratively.
- Highlight contextual details and build visual indicators.
- Eliminate financial incentives.
- Crack down on computational amplification.
- Adequately moderate non-English content.
- Pay attention to audio/visual forms of mis- and dis-information.
- Provide metadata to trusted partners.
- Build fact-checking and verification tools.
- Build "authenticity engines"
- Work on solutions specifically aimed at minimizing the impact of filter bubbles: a. Let users customize feed and search algorithms. b. Diversify exposure to different people and views. c. Allow users to consume information privately. d. Change the terminology used by the social networks.

What could national governments do?

- Commission research to map information disorder.
- Regulate ad networks.
- Require transparency around Facebook ads.
- Support public service media organizations and local news outlets.
- Roll out advanced cybersecurity training.
- Enforce minimum levels of public service news on to the platforms.

What could media organizations do?

- Collaborate.
- Agree policies on strategic silence.
- Ensure strong ethical standards across all media.
- Debunk sources as well as content.
- Produce more news literacy segments and features.
- Tell stories about the scale and threat posed by information disorder.
- Focus on improving the quality of headlines.
- Do not disseminate fabricated content.

What could civil society do?

- Educate the public about the threat of information disorder.
- Act as honest brokers.

What could education ministries do?

- Work internationally to create a standardized news literacy curriculum.
- Work with libraries.
- Update journalism school curricula.

What could funding bodies do?

- Provide support for testing solutions.
- Support technological solutions.
- Support programs teaching people critical research and information skills.

What civil society and media can do

Civil society through its groups, networks, and citizens acting on their own provides an important constituency in addressing mis/disinformation. And yet civil society's critical role is sometimes not adequately supported. As Dr. Joan Donovan, director and lead researcher of the Technology and Social Change Research project at the Shorenstein Center at Harvard Kennedy School, has said, "The lack of attention to civil society responses is a major gap in the research and it is becoming increasingly clear that the guidance for journalists does not translate easily to civil society."[173] In order to activate citizen approaches, Donovan advocates a three-pronged approach for strengthening civil society to counter disinformation: Detect, Document, and Debunk (*aka* the 3Ds approach). The 3Ds approach, however, takes time and a lot of research, planning, and strategy. The guidance and examples in Annex 5, Quick resources for Planning a Disinformation Strategy, may help with planning and design for civil society-led approaches to countering and preventing mis/disinformation.

Supply-side responses (producers)

Many approaches seek to address the "supply side" of disinformation and misinformation, attempting to limit the influx or supply of false or misleading information to media systems and publics.[174] The following 10 supply-side responses (fact-checking, redirecting, debunking, pre-bunking, legal/policy, counter-disinformation campaigns, open-source intelligence, supporting local journalism support, and increasing transparency in journalism and advertiser outreach) are among the emerging practices that provide examples of how civil society groups and individual citizens are countering dis/misinformation by targeting it where and how it originates and spreads today.

Why does fact-checking matter, and is it effective? Fact-checking, just like journalism, is about informing audiences about the truth. Even if it is a limited audience, or if people disagree, calling out disinformation attempts to hold people and institutions accountable. Sometimes fact-checking efforts are picked up by larger outlets and reach wider audiences, but even if they are, publishing fact-checking is only the first step in an incremental process. When watchdogs expose untruths, this can become a resource for public action, particularly when mobilized by political campaigns or social movements. They can also identify trends and help trigger an institutional response by regulators, courts, or legislators.

In the last few years, fact-checking organizations have become more effective and grown tremendously. The Reporter's Lab hosts a database of reporting projects that regularly debunk political misinformation and viral hoaxes. As of 2019, the database counts 210 active fact-checking organizations in 16 countries, up from 59 sites tallied in 2014.[175]

Bringing these organizations together, the International Fact Checking Network provides resources, monitors trends, and promotes basic standards for fact-checking organizations through its code of principles.[176] Many of these organizations work in USAID program countries; it is worth reaching out to them when developing disinformation programming. A few examples of fact-checking organizations are included in Annex 3: Emerging Solutions.

The Effectiveness of Fact-Checking

The effectiveness in fact-checking is the subject of numerous studies. These studies have often produced contradictory results. In one key study, researchers from three universities teamed up to analyze the results of the studies and determine the effectiveness of fact-checking across research studies. The resulting study helps shed some light on the fact-checking landscape and the ins and outs of what works in fact checking. It found that:

- Fact-checking with graphical elements is less effective than those without,[177] and simple messages are more effective when informing the public on false information.[178]
- Fact-checking an entire statement rather than individual elements is more effective.
- When fact-checking, debunking an idea is more effective when it is refuting ideas in line, as opposed to, a person's own ideologies.[179]

While it is important to consider the limitations of fact-checking—it is not a panacea for all disinformation—fact-checking is highly prized by many people and is especially important for media outlets and others who are in the business of investigating and reporting facts.

#2: REDIRECTION METHOD

The Redirect Method primarily relies on advertising using an online advertising platform such as Google AdWords, targeting tools and algorithms to combat online radicalization that comes from the spread and threat of dangerous, misleading information.

The method redirects users through ads who seek to access mis/disinformation online to curated YouTube videos uploaded by individuals around the world that debunk these posts, videos, or website messages. The Redirect Method, a method used to target individuals susceptible to ISIS radicalization via recruiting measures, is being adapted in several countries to combat vaccine hesitancy and hate speech.[180] The method was developed by a

countries to combat vaccine hesitancy and hate speech.[180] The method was developed by a collaboration among private and civil society organizations and is documented on The Redirect Method website. The collaboration provides 44 steps in the organization of an online redirect campaign. Other examples of the use of redirection are being employed by groups around the world.

#3: DEBUNKING AND DISCREDITING

The Global Engagement Center (GEC) at the U.S. Department of State recommends a combined debunking and discrediting approach, which is explained in GEC Counter-Disinformation Dispatches #2: Three Ways to Counter Disinformation and GEC Counter-Disinformation Dispatches #4: What Works in Debunking.

Key points include:

- A "counter-brand" approach, which involves discrediting the "brand"—the credibility and reputation—of those making false allegations.
- Highlighting false claims seen as obviously absurd and particularly offensive and objectionable by target audiences
- Changing the frame from the false allegation to the misdeeds and lack of credibility of those spreading disinformation
- Creating moral outrage: truth is a sacred value; spreading vicious disinformation violates this sacred value, creating moral outrage.
- Recognizing that the mind often reasons by associations rather than logic.

See the GEC Counter-Disinformation Dispatches for more details.

#4: PREBUNKING

As a measure to counter disinformation and make debunking more impactful, Donovan recommends prebunking, which she defines as "an offensive strategy that refers to anticipating what disinformation is likely to be repeated by politicians, pundits, and provocateurs during key events and having already prepared a response based on past fact checks."[181] Prebunking is drawn from inoculation theory, which seeks to explain how an attitude or belief can be protected against persuasion, and people can build up an immunity to mis/disinformation.[182]

Five Steps to Execute a Prebunking Strategy

1) Take a look at fact-checking websites and databases to get a sense of the trends in misinformation.

2) Map out which misinformation trends are popular on Twitter (or other social media) in politicians' stump speeches.

3) Find additional source material with the facts about the misinformation

Other researchers have also pursued the prebunking track, including Dutch researchers who developed a game called Bad News,[183] which helps people spot misinformation and disinformation. According to the developers, it helps people to talk about the truth and reduces their susceptibility to misinformation. Based on initial research of users who have played Bad News, it has been an effective approach and led to improved psychological immunity against online disinformation. Another prebunking game (funded by the Global Engagement Center) is called Harmony Square. The theory behind this game also draws on inoculation theory; learn more about this in this Harvard Misinformation Review article.

likely to be repeated.

4) Prepare your social networks for the high potential for misinformation.

5) Turn "prebunk into debunk" by immediately posting correct information anywhere you can. Finally, in using prebunking techniques, she counsels that speed matters.

Source: Dr. Joan Donovan, Shorenstein Center, Harvard, [@BostonJoan]. Tweets. https://twitter.com/BostonJoan

#5: Legal and Policymaking Advocacy

As information pollution and disinformation begin to affect the ways in which democracies function, governments have become aware of the importance of combating false information online. While there is government support to find ways to combat the information disorder, democracies are struggling to find the best way to regulate disinformation online. There is an inherent contradiction between the democratic ideal of free speech and the regulation of online content. Despite this, some governments have been developing policies that are intended to combat the spread of hate speech and disinformation, while simultaneously working hard to preserve free speech online.

However, when democracies pass new laws to curb the spread of disinformation, authoritarians can adopt these laws as yet another tool to criminalize free expression. So, there is great risk of making the situation worse by regulating online speech. A very clear example is Germany's NetzDG law (requiring social media to take down harmful content) that was later adopted by Russia, Venezuela, and other countries as a means of silencing opposition.[184]

The Advisory Network to the Media Freedom Coalition, a group of 17 national, regional, and international organizations, delivered a statement at the ministerial meeting of the 2020 Global Conference for Media Freedom and put forward the following guidance to support legal and policymaking advocacy that seeks to deal with the disinformation problem:

- Respect media freedom while tackling disinformation.
- Efforts to combat disinformation and "fake news" must start with governments, which should not criminalize this but also must commit to not perpetuating disinformation and fake news.
- Counter the criminalization of journalism through so-called anti-fake news laws and anti-terrorism laws.
- Promote media engagement in countering disinformation by expanding access to information mechanisms and by supporting journalistic investigations revealing the

sources and dissemination patterns of disinformation and highlighting the role of government representatives in spreading disinformation.[185]

Because disinformation is considered a wicked problem, some have called for new ways to regulate the free flow of information with co-regulatory models. Some ideas were outlined recently in the Disinformation as a wicked problem: Why we need co-regulatory frameworks policy-paper from Brookings Institution.[186]

> To effectively manage disinformation and related online problems, governments and platforms will need to develop an architecture to promote collaboration and build trust among stakeholders. There are several models for multi-stakeholder collaboration, among them the industry-led Global Internet Forum to Counter Terrorism (GIFCT) and the government-led Information Sharing and Analysis Centers (ISACs). Those that prove successful have in common continuous adaptation and innovation and a focus on trust-building and information-sharing.

A few examples of resources for legal and policy approaches are included in Annex 3: Emerging Solutions.

#7: MESSAGING CAMPAIGNS

Messaging campaigns have provided a new front where disinformation can easily spread to large amounts of people. Messaging campaigns refer to messages that contain false information that pass through networks like Facebook Messenger, Twitter Direct Message, and WhatsApp private and group messaging. These large-scale campaigns are difficult to track because of the encryption security available on most messaging apps. Despite this increased difficulty, there are new top-down and bottom-up approaches that are being utilized to help stop the prevalence of messaging disinformation campaigns.

A few examples of resources counter-disinformation campaigns are included in Annex 3: Emerging Solutions.

#8 LOCAL JOURNALISM SUPPORT

Bolstering local journalism is key to countering and preventing disinformation. Two pivotal studies provide an evidence-based set of findings that show why support to quality, independent media is key. One study notes that "results indicate local newspapers hold their governments accountable, keeping municipal borrowing costs low and ultimately saving local taxpayers money."[187] A second study observed, "We find newspapers have a robust positive effect on political participation, with one additional newspaper increasing both presidential and congressional turnout by approximately 0.3 percentage points."[188]Furthermore, major research undertaken by the United Nations and the International Center for Journalists cited quality journalism as a major force for identifying and exposing disinformation,[189] citing that the 'viral load' of disinformation will only grow if journalism continues to suffer death blows inflicted by the (COVID-19) pandemic. The importance of good, quality local journalism as the key to fighting false news and noxious content was nicely argued in a pair of articles written by Emily Bell at the Tow Center for Digital Journalism at Columbia University's

Graduate School of Journalism and a Guardian columnist and by Alan Rusbridger former editor of the Guardian and Chair of the Reuters Institute for the Study of Journalism. As Rusbridger sums it up "against evasion and lies, good journalism is all we have."[190]

As social media has expanded as a primary way for many individuals to receive their news, there has been a decrease in funding for local journalism and a need for better support.[191] This change is largely a result of advertising revenue that is being redirected towards online media sources. Total global newspaper advertising revenue may lose about $23.8 billion in annual revenues from 2012 to 2021. More than 10 percent of this decline, around $3 billion, is an estimated loss of annual revenue for local news media around the world.[192] Television news is still holding its own in places such as the Philippines; however, more than 80 percent of Filipinos say they now go online for their news and spend four of their 10 hours online accessing social media.[193]

In the vacuum created by the loss of local newspapers, the Brookings Institution finds that readers turned to outlets covering national stories that may have strong partisan leanings or concentrate on partisan conflict.[194] As a result, Brookings observed that "in places where news consumers cannot balance their news diet with local alternatives, voters tend to be more politically polarized."[195] Understanding the value of local journalism and the challenges it faces, some civil society organizations have committed themselves to revitalizing and protecting local news efforts. Through providing funding, training, and other resources to journalists and local outlets, organizations ensure that communities have sustained access to information through a more local lens and without a dependence on large-scale, potentially partisan media houses.

Photo: IREX Learn to Discern (L2D) media literacy program in Jordan

Some prominent international organizations acting in this space include:

Internews builds lasting change by ensuring people have access to quality, local information. To do so, Internews works with local partners to grow sustainable organizations and offers capacity-building programs for media professions, human rights activists, and information entrepreneurs.[196]

[redacted]

IREX promotes "vibrant information and media systems." IREX supports journalism and media organizations through trainings on reporting, media law, media safety, and digital security. IREX also provides additional support to consumers via media literacy programs, training citizen journalists, and diversifying and distributing television content.[197]

[redacted]

International Center for Journalists (ICFJ) seeks to build the expertise and storytelling skills of journalists around the world. ICFJ focuses on five key areas: news innovation, investigative reporting, global exchange programs, specialty journalism, and diversity promotion.[198]

Over the years, media development civil society organizations have been formed and have become active in most parts of the world. A useful list of local, regional, and international organizations can be found at the Global Forum for Media Development.

#9 TRANSPARENCY IN JOURNALISM

As the spread of disinformation online grows and even reputable news agencies have mistakenly shared false information, there is a need for better trust and transparency standards to hold media agencies accountable. Some very basic standards such as bylined articles, public display of address and registration information, and disclosure of funding sources and editorial board, for example, can assist readers to understand the sources and reliability of the information they read. By developing standards for media agencies, journalism will be held to greater account for what it publishes. Examples of ongoing initiatives aimed at rebuilding trust and strengthening transparency in journalism are included in Annex 3: Emerging Solutions.

#10 ADVERTISER OUTREACH

In order to disrupt the funding and financial incentive to disinform, attention has also turned to the advertising industry, particularly with online advertising. A good example of this is the concerted response to the discovery of the websites traced to the village of Veles outside of Skopje, Macedonia, which showed how easy it was to exploit the digital advertising model to flood the news ecosystem with fabricated content. *Wired Magazine* profiled the Veles case study in 2017.[199]

As most online advertisers are unaware of the disinformation risk of the domains featuring their ads due to the automated process of ad placement, they inadvertently are funding and amplifying platforms that disinform.[200] Thus, cutting this financial support found in the ad-tech space would obstruct disinformation actors from spreading messaging online. Efforts have been made to inform advertisers of their risks, such as the threat to brand safety by being placed next to objectionable content, through conducting research and assessments of online media content. Additionally, with this data, organizations hope to aim to redirect funding to higher-quality news domains, improve regulatory and market environments, and support innovative and sustainable models for increasing revenues and reach.

A few examples of advertiser outreach are included in Annex 3: Emerging Solutions.

Demand-Side Solutions

Other groups have centered their efforts on targeting the opposite side of disinformation, the "demand." With demand-side solutions, the focus shifts to reducing the general societal acceptance or tolerance of distorted and fabricated items.[201] These approaches are often viewed as more important for sustaining long-term impacts as they concentrate on why and how disinformation influences society and the strategies to develop societal resilience against it. Different demand-side solutions often look to governments, tech giants, and social media companies to promote education and training, raise awareness, or advance other efforts toward building more resilience against disinformation.

MEDIA LITERACY, EDUCATION & TRAINING

In response to today's challenging information environment, many civil society organizations view media literacy, education and training programs as critical mechanisms for hardening communities against disinformation. These projects strive to assist consumers in understanding how to access, assess, and critically evaluate different types of media, as well as process and analyze media messages. They seek to promote the development of healthy habits and behaviors for information consumers and provide them with the tools and other resources for identifying accurate, high-quality content from disinformation.

A few examples of media literacy, education and training are included in Annex 3: Emerging Solutions.

CIVIL SOCIETY ADVOCACY AND PUBLIC AWARENESS CAMPAIGNS

A multitude of civil society organizations, both local and international, are pushing for better accountability and transparency. Other approaches are concentrating on providing research and advocacy on the vulnerability of countries to disinformation and exposure.

To build societal resilience against disinformation, the public must be aware of and understand the issue. Public awareness campaigns serve to inform and educate communities about disinformation, build public recognition of it, and promote actions or other resources for combating it.

A few examples of civil society advocacy and public awareness campaigns included in Annex 3: Emerging Solutions.

DIGITAL LITERACY & SECURITY

Civil society organizations have also sought to increase digital literacy and security skills to empower the online community to identify and counter disinformation. Digital literacy trainings provide lessons and resources for understanding how to identify, evaluate, and compose information on digital platforms. This type of literacy expands understanding of how to navigate, read, and interpret in a digital environment, as opposed to media literacy that focuses primarily on understanding and interpreting journalism. Furthermore, digital security initiatives produce and provide tools and behaviors for maintaining safety of personal identities and information online that protect people and groups from becoming the target of mis/disinformation.

An example is the Digital Society Project, which seeks to provide data on the intersection of politics and social media. Topics include online censorship, polarization, misinformation campaigns, coordinated information operations, and foreign influence in and monitoring of domestic politics. It uses the Varieties of Democracy (V-Dem) framework, also used by USAID in Journey to Self-Reliance (JSR) metrics, to assess various digital issues including misinformation.

Elections-focused programming efforts (both supply- and demand-side solution)

To defend the integrity of democratic elections worldwide, reducing disinformation has been central to elections-focused programming in all these common goals:

- Building the capacity of election management bodies to combat disinformation during electoral periods.
- Strengthening and adapting citizen election observation to incorporate disinformation monitoring and exposure efforts.
- Adapting international election observation to integrate monitoring of the information environment, including social media.
- Supporting independent media election coverage capacity and resources to increase the supply of accurate information during elections.
- Increasing incentives for ethical online conduct by political parties and campaigns (such as facilitating norms for online Codes of Ethics for political parties).

As social media has changed how audiences consume and spread political information, focus has shifted to monitoring the transformed information environment before and during election periods. Emphasis has been placed on citizen election and campaign finance monitoring, focusing on exposing disinformation on social media and other digital platforms during election campaigns. Election management bodies (EMBs) tend to have significant power to regulate speech and can set some ground rules. Media platforms have been compelled by EMBs to be transparent about political advertising during an election season.

A report produced by the Kofi Annan Commission on Elections and Democracy in the Digital Age puts forward recommendations to governments and internet platforms to safeguard the legitimacy of elections. As the report notes, "For the foreseeable future, elections in the democracies of the Global South will be focal points for networked hate speech, disinformation, external interference, and domestic manipulation."[202]

Disinformation assessments have also been incorporated within international election observation and management missions. With this, organizations have offered trainings and guidance for election bodies and other stakeholders in regulating speech, campaign violations, and the prosecution of violations.

A few examples of elections-focused programming and disinformation assessments are included in Annex 3: Emerging Solutions.

There is also advocacy working to motivate social networking platforms, advertisers, governments, and other parties to improve the information environment with pertinence to the integrity of elections. In particular, advocacy aimed to achieve greater transparency in political advertising, close fake accounts and bots, and de-monetize suppliers of

political advertising, close fake accounts and bots, and de-monetize suppliers of disinformation. An example of this is European Commission's "Code of Practice on Disinformation," a self-regulatory Code of Practice to fight disinformation and ensure transparent, fair and trustworthy online campaign activities ahead of the 2019 European elections. Signatories included Facebook, Google, Twitter, Mozilla, Microsoft, and TikTok.[203]

In more closed environments, with governments seeking to suppress opposition, trainings for political candidates, parties, and activists have become valuable resources for learning how to disseminate messages in adverse and media-saturated communities. An example of this work in more closed environments is NDI's "Raising Voices in Closing Spaces," a workbook offering a step-by-step approach to strategic communications planning for citizen election observation groups and other civil society organizations to break through in closing or closed political environments. The guide offers strategies, tactics, and hypothetical and real-world examples for overcoming challenges in repressive environments.[204]

A. What governments can do

Governments allocate state resources and execute strategies in order to protect their citizens and the integrity of their information systems/democracy from the harms imposed by information disorder. This proves difficult to accomplish through traditional, analog legislative action because disinformation networks transcend the national borders where legislative mandates are confined. We will examine below several countries that have attempted to quell the production and dissemination of disinformation through legal and regulatory measures.

Though it can be well-intentioned, criminalizing the spread of disinformation in digital spaces presents even more negative ramifications for democratic governance. In countries where disinformation legislation has been introduced, creating or sharing digital content and information becomes risky and punishable. The resulting chilling effect ultimately hampers democratic protections for a free press and the freedom of expression. Likewise, any whole scale internet or platform shutdowns pose the same anti-democratic results while also creating information vacuums within communities amid public health or safety crises. As USAID negotiates the entangled, disinformation landscapes that intersect global and local programming, programs should be wary both of legislative efforts to limit expression online and internet/platform blackouts for the dangers they pose to democracy.

1. Whole-of-Government Approaches

There are very few examples of whole-of-government approaches (integrated activities of government agencies increasing coherence in policy development on an issue) in the disinformation space. Policymaking is likely to take time and will need to be significantly informed by the private sector and civil society. Two examples in Scandinavia are discussed to illustrate this approach.

Some governments have seen success in whole-of-government approaches in which government agencies, social institutions, and civil society collaborate to collectively close the digital, education, and social fissures that disinformation campaigns need to infiltrate media and political landscapes.

In Sweden, for example, the whole-of-government approach (a comprehensive approach that involves government, private and civil society sectors and media) attempts to be as multi-faceted as the issue itself. Focused on maintaining the integrity of its electoral processes, the Swedish government has poured resources into shoring up its electoral infrastructure by safeguarding electronic voting machines and providing new training for civil servants and voting officials. The government, electoral bureaus and cybersecurity agencies have committed to forming an inter-agency cooperative to combat the issue.[205] Sweden has also focused on bettering public media literacy through enhancing high school curriculums as a long-term strategy against information disorder. Lastly, five of the country's most prominent media news outlets have created a scrupulous and highly visible fact-checking collaboration to maintain the integrity of Swedish media and spread mutual oversight over multiple media stakeholders.[206] Sweden's approach has required weaving together

resources to defend the many susceptible ports of entry that misinformation and disinformation penetrate in media landscapes and societies.

In Finland, response to the global information crisis has included a sweeping media literacy education effort. In junior high and high schools, Finnish students use a digital literacy toolkit in order to understand how to implement best practices in their own social media engagement. In the world of electoral and media professions, Finland has dedicated resources to disinformation trainings and workshops for civil servants and journalists.[207] This approach seeks use of public pedagogy and a more robust media literacy secondary curriculum to intervene at every level of societal strata, in hopes that many small contributions by individuals will add up to big structural changes that make a nation less vulnerable to disinformation. Critics of the Finnish strategy call it too slow moving to address such a fast-moving problem and worry that there are not concrete ways to measure its success.[208] Taiwan and Lithuania provide other good examples of whole-of-government approaches to mis/disinformation.[209]

2. Strategic communications and State coalitions

On a broader scale than the individual efforts of governments within their respective countries, coalitions among states and policy stakeholders are expanding global efforts to fight disinformation, such as EUvsDisinfo and the Atlantic Council.

The EUvsDisinfo is a flagship strategic communications project of the European External Action Services, East StratCom Taskforce since 2015 to "increase public awareness and understanding of the Kremlin's disinformation operations, and to help citizens in Europe and beyond develop resistance to digital disinformation and media manipulation."[210] EUvsDisinfo has hands in three major areas to stymie global disinformation. The project conducts data analytics of media spaces to identify and publicize disinformation campaigns started by the Kremlin or pro-Kremlin media outlets; the data is archived in an open-source, dedicated database that tracks disinformation around the world. The project also publishes summaries, articles, and reports both within the research community and for wider general readership in

order to make a technical and complex issue accessible to the broader global communities it impacts. Lastly, EUvsDisinfo engages with governments, media outlets, universities, and civil society organizations by offering training resources to better prepare and problem-solve around issues of information disorder.

Another example of strategic communications is The Atlantic Council, a "nonpartisan organization that galvanizes U.S. leadership and engagement in the world, with allies and partners, to shape solutions to global challenges."[211] Among the numerous capacities through which Atlantic Council conducts international work, their recently published *Democratic Defense against Disinformation 2.0* provides a current snapshot of global disinformation campaigns at work, including proposed solutions, progress made, and criticisms and recommendations made to the U.S. executive and judiciary branches regarding urgent steps needed in the path toward a healthier information ecosystem.

3. Legal and regulatory measures

While there are examples of legal and regulatory measures intended to address disinformation, the difficulty is that many are controversial because they also restrict freedom of expression.

To achieve good legal and regulatory approaches to disinformation will require civil society and independent media organizations to have a significant voice in policymaking through multi-stakeholder collaboration. Governments face very real tradeoffs in developing regulations on digital rights, disinformation and dangerous speech, while balancing protections for freedom of speech, transparency, and privacy. Regime type is an important variable to consider.

Guidance on good practices regarding regulation and policymaking in this area was put forward by the UN Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the OAS Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information, with Article 19 and the Center for Law and Democracy. These standards can be useful in evaluating legal and policymaking processes and frameworks; some important ones are excerpted here from the section on Standards on Disinformation and Propaganda:

> a) General prohibitions on the dissemination of information based on vague and ambiguous ideas, including "false news" or "non-objective information," are incompatible with international standards for restrictions on freedom of expression, as set out in paragraph 1(a), and should be abolished.
>
> b) Criminal defamation laws are unduly restrictive and should be abolished. Civil law rules on liability for false and defamatory statements are legitimate only if defendants are given a full opportunity and fail to prove the truth of those statements and benefit from other defenses, such as fair comment.

c) State actors should not make, sponsor, encourage, or further disseminate statements which they know or reasonably should know to be false (disinformation) or which demonstrate a reckless disregard for verifiable information (propaganda).

d) State actors should, in accordance with their domestic and international legal obligations and their public duties, take care to ensure that they disseminate reliable and trustworthy information, including about matters of public interest, such as the economy, public health, security, and the environment.[212]

B. WHAT SOCIAL MEDIA PLATFORMS CAN DO

The issue of misinformation reminds many of the saying, "A lie can travel halfway around the world while the truth is putting on its shoes."[213] This saying becomes especially concerning in reference to social media platforms, whose technologies have significantly impacted the ways in which people communicate and share information. Because social media offers free, open access to reach widespread global publics, discussions have turned to calling for the platforms to respond to the online environments and issues they have created. More specifically, focus has narrowed in on social media's automated algorithms that have provided an easy and rapid mechanism for misinformation to spread and be amplified, leaving truth to be buried amongst this free reach of falsehoods.[214] Thus, critics argue that social media platforms must take greater responsibility for themselves and increase platform regulation and oversight and question whether self-regulation is sufficient.

With the algorithms of Facebook, Google, Twitter, and the other social media platforms driving how people access and consume today's information, these organizations hold the power to change the way information flows online and to address the conditions they created

for enabling misinformation to flourish.[215] Social media companies' business models amplify and further complicate the existing issues of misinformation, disinformation, and malinformation. Although these companies are part of this problem, they can and will need to be collaborators for determining solutions and best practices.

"It is difficult for the truth to get out if someone shares falsehoods and benefits from free reach and automated algorithmic amplification. The facts must be given a chance for a proportional response."

—Joshua Lowcock, chief digital and global brand safety officer at Universal McCann

What is platform accountability?

As explained in previous sections of the primer, the ability for forms of disinformation and misinformation to spread rapidly via social media positions social media companies not only as gatekeepers and facilitators of information-sharing and communication, but also for political interaction and engagement. Thus, policymakers, civil society, academia, and the public have become increasingly interested in understanding the political implications of social media, calling for operations that would make the platforms more democratically accountable.[215] Tactics to expose disinformation—such as: 1) fact-checking, 2) content moderation, and 3) labeling of sources—can be found in the next subsection. Accountability

moderation, and 3) labeling of sources—can be found in the next subsection. Accountability initiatives have been growing in response to increased awareness of disinformation: for example, see Facebook's Community Standards Enforcement Report.

In considering how and by whom platforms should be held accountable, discussion extends to platform governance. Robert Gorwa, a scholar from the University of Oxford, explains platform governance as an "approach necessitating an understanding of technical systems (platforms) and an appreciation for the inherently global arena within which these platform companies' function."[217] In essence, the approach necessitates more knowledge about how platforms govern or influence publics via their platform practices, policies, and affordances while also acknowledging that their conduct is informed by local, national, or international forms of governance.[218]

Currently, there are three governance "lenses" being utilized for thinking about social media platform regulation and policy: 1) self-regulation, 2) external governance, and 3) co-governance. The first lens represents the presently dominant mode of governance, a rather laissez-faire relationship between governing institutions and social media companies, in which the companies dictate their own platform improvements and transparency efforts. The second lens encapsulates governments enacting legislation regulating social media platforms. Lastly, the third lens seeks to provide greater democratic accountability without severe disruption to the status quo. As noted earlier, the challenge with this approach is striking a balance between combatting the spread of disinformation and harmful content while also preserving free expression online. Proponents of this lens have proposed the creation of specialty organizations, third-party adjudication systems, or ombudsman programs that would act to address and investigate platform complaints as well as set ethical frameworks and standards, among other tasks.[219] Some examples are ICANN, a not-for-profit, public benefit organization that regulates domain names, and the Internet Governance Forum, which includes all stakeholders on governance questions in open and inclusive fora.

Current platform initiatives

As a result of public concern, this has become an exciting and rapidly moving space where platforms are developing solutions to be responsible to their consumers, as well as where government and civil society organizations are leading additional efforts to influence platform initiatives. There are many discussions about the unregulated power of social media and other digital platforms. Below is a non-exhaustive list of how social media companies are

creating solutions internally and how global civil society groups and human rights defenders are encouraging greater platform accountability.

Models for Platform Governance: An essay series on global platforms' emerging economic and social power, Centre for International Governance Innovation

Content Regulation and Human Rights, Global Network Initiative

The following table highlights a few social media initiatives (as of late 2020) to address disinformation and misinformation.

Initiative	Social Media Platform	Notes

Fact-Checking Program	Facebook	With a three-step approach of "identify, review and act," Facebook has independent fact-checkers assess and rate the accuracy of stories through original research. Facebook may act by reducing distribution, including sharing warnings, notifying previous sharers of misinformation, applying misinformation labels, and removing incentives for repeat offenders.[220]
Oversight Board	Facebook	This internationally diverse group formed to protect freedom of expression by making independent decisions on content issues and providing policy advisory opinions to Facebook.[221] *Note:* Facebook's Oversight Board has been the subject of much criticism, as highlighted in an article in the Harvard Business Review.[222]
Expansion of Hate Speech Policy to Include Holocaust Denial	Facebook	In response to the rise in anti-Semitism globally, Facebook expanded its hate speech policy to include Holocaust denial. Any content that "denies or distorts the Holocaust" is now banned. *Note:* Facebook's decision to undertake action on hate speech related to Holocaust denial is also perceived with skepticism and critique.
New Warning Labels about COVID-19 Information	Twitter	Twitter added new labels on tweets about COVID-19 that link to a Twitter-curated page or external trusted source for additional information on the claims of the tweet. Tweets conflicting with public health experts would be covered by a warning from Twitter. Misleading information with a severe propensity for harm would be removed from the site.[223]
Birdwatch	Twitter	Twitter is potentially developing a new product called "Birdwatch," which will likely involve crowdsourcing to address disinformation and provide more context for tweets in the form of notes.
Encouraging Users to Read Articles Before Sharing	Twitter	Twitter tested a feature on Android that will prompt a message suggesting the user should read an article before sharing it.[224]

* Regarding Facebook, numerous experts from academia, journalism, civil society, government, and other sectors point out the outsize problem that the company creates in terms of the spread of disinformation. These critiques make clear that the social media giant struggles to strike a balance between allowing free expression and enabling the spread of potentially dangerous hate speech, extremism and disinformation.[225] These concerns should be taken very seriously, as Facebook is the biggest social network in the world, with more than 2.7 billion users.[226]

VII. PART SEVEN: TEN THINGS USAID AND ITS PARTNERS

VII. PART SEVEN: TEN THINGS USAID AND ITS PARTNERS CAN DO TO COUNTER AND PREVENT DISINFORMATION

As noted in the primer's introduction, information disorder is a wicked problem. Like other seemingly insurmountable challenges, there is a need to tackle the problem in a number of ways. Any solutions to the root causes of information disorder will need to work in concert with other issues, such as better governance, cleaner elections, stronger accountability—the whole spectrum of DRG issues. Simply put, disinformation is the big spoiler of all the DRG issues. It's not just about doing one little thing, however; it's about a lot of little things, done in an interactive, adaptive way and over a period of time.

Photo: ©2020 Shutterstock/Chanonnat Srisura

While it may not be possible to eradicate information disorder, **doing nothing is not an option**. There are steps that must be taken to help alleviate the problem and ensure that the progress and successes USAID has made around the world are not jeopardized.

Note: you will want to make your work as context specific as possible and do original, diagnostic research to determine the best course of action for implementing more effective strategies.

A. TEN STEPS FOR DEALING WITH DISINFORMATION

1) Conduct a disinformation diagnostic. Take stock of the information environment in your country. The Public Affairs Section, the Bureau of Intelligence and Research, and/or the Global Engagement Center at the State Department may have some information, often on a smaller scale. The assessment could be part of a standalone media program or integrated into larger governance programming. An assessment should understand cultural notions attached to information legitimacy and norms for accepting the veracity of information. Some regions are culturally preconditioned to be much more skeptical (e.g., former Soviet republics) while others experience wide deference to authorities (many Asian societies).

An assessment would look for answers to the following questions.

a. What are the biggest threats/vulnerabilities for disinformation?

a. What are likely disinformation themes, aims, and main sources of disinformation?

b. Who is most vulnerable to the disinformation being spread?

c. What are the common disinformation narratives?

d. Who are the main disinformation purveyors?

e. How does it spread across communities, and who are the most vulnerable?

b. What is the media ecosystem in your country like?

a. What is the makeup of mainstream media, including state, public service, and private media as well as community media and alternative media?

b. What is the social media landscape like in your country, and what are the most popular platforms (i.e., Facebook, Twitter, WhatsApp, etc.)?

c. What are the most common ways that different segments of the population consume information?

d. Where do different people get their information? Is it online or through TV, radio, print or peer-to-peer sources?

e. How do people access the internet? (mobile/desktop)

c. What is the media literacy landscape? What are the media literacy levels?

d. What is the rate of mobile and internet penetration?

e. What does current research say about the state and nature of the media and social media, including levels of press freedom, internet freedom and any threats or vulnerabilities that exist regarding access to information?[227]

2) **Carry out actor mapping.** Assess the key stakeholders who contribute in positive and negative ways to the media and information ecosystem in your country.

a. Who are the malign actors?

b. Are there civil society organizations, associations, or institutions that seek to root out disinformation, increase public awareness, engage in fact checking, or contribute in a positive way to countering the disinformation problem?

c. Who are the allies, and who are the foes? Are there media, civil society, academics, think tanks, businesses, or government agencies that you can work with?

d. Are there gaps in the actors (fact-checking or media development organizations, for example), and could these gaps be addressed by partnering with international CSOs or companies?

3) **Support media literacy initiatives.** Given the changing nature of mass media and news and information consumption habits, youth as well as other segments of the population benefit from opportunities to learn critical thinking skills and to acquire the knowledge to spot disinformation and other aspects of information disorder. Participants of media literacy programming can also serve as good boosters or spreaders of critical thinking.

However, media literacy, like other aspects of knowledge and learning, requires "booster shots" in which individuals learning about how the system works can get updated skills and knowledge on new developments and techniques of media manipulation. Efforts to get media literacy into the curriculum of primary and secondary schools on a broad scale are considered best practice, but media literacy programs can be run through schools, libraries, and other community-level partnerships, and they should be integrated across subjects, as well as serve as part of a civic education curriculum. If possible, they should also target a variety of age groups, from children to adults and seniors. As part of media literacy efforts, include ideas and tactics for prebunking.

4) **Fund independent media and local public interest journalism.** High-quality, independent, local news is under threat. Endangered by changing business models, big tech's domination of the digital advertising market, and pressures stemming from media capture, concerns over press freedom, and the overall safety of journalists, the future of news will require substantial investment and partnership with donors as well as experimentation with different models and approaches. As advocated by the Knight Foundation, one of the best forms of resilience against mis- and disinformation is journalism.[228]

5) **Support media monitoring and fact-checking initiatives specifically.** As shared by Lithuania's Elves in their fight against the disinformation trolls, you cannot fight disinformation with disinformation. You need to fight back with the truth and good reporting. Local actors also need to know where and how disinformation is spreading. This same sentiment is shared by leading journalists and experts on disinformation, who argue that the best defense against the lies, propaganda, conspiracy theories, and other bad information that is circulating is the truth.[229] Truth, transparency, and facts are an essential piece of fighting back against the wicked problem of information disorder. Distortion of the facts is a moving target, so the ability to do fact-checking needs to be sustained over time.

6) **Stay up to speed.** Disinformation is a growing field and threat, and there is a need to stay current on new research and tools for identifying disinformation trends and threats and to stay abreast of the different disinformation narratives that could undermine USAID's work. While this requires sustained effort, the cost of inattention is too high if ignored.

7) **Support internet governance and digital rights initiatives.** Digitalization and the spread of internet technologies has already led to the next frontier of information access. It has, unfortunately, also provided another way to prevent citizens from accessing information. Criminals and politicians are known to muddy the waters and engage in malign activities, often using the internet and the disinformation playbook to gain money and power. Just as human rights advocates have argued that internet access is a human right and that there is a fundamental right to access of information, there is also a right *not* to be disinformed. The internet freedom and digital rights sector that USAID has supported and partnered with around the world is a vital source of expertise and collaboration in the fight against disinformation. Consider potential partnerships with digital security trainers, digital forensic investigators, digital literacy specialists, internet governance experts, and others who can work with your local USAID programs as part of democracy and governance programs. For example, working with the Internews-led

Internet Freedom Consortium, a program funded by USAID, can help to address these issues in countries where USAID works.[230]

8) **Engage government.** Where it is possible, reach out to local government and national government officials to begin a dialogue, if one is not already in place, about how USAID can partner in countering and preventing disinformation. The information disorder problem puts USAID's work at risk, but it also tarnishes local economies, jeopardizes public health systems, and wreaks havoc on peace and stability. USAID can be a positive partner in putting out strategic messaging and countering disinformation narratives that may occur in a given country. USAID can also partner with ministries and schools to support media literacy efforts or broker partnerships to help support investments in needed infrastructure or systems required to counter and prevent disinformation.

9) **Collaborate and engage with other international partners** (civil society, media, social media platforms, internet governance forums, and other donors). Countering and preventing disinformation programming is a growing field and one that is likely to be a major component of programs for some time to come. Given the rising influence of both Chinese- and Russian-backed disinformation efforts that put USAID's democracy and governance programs at risk, look for ways to form synergies with others working to support and strengthen democracy in the country where you operate. Collaborative partnership will be stronger, and funding can be better leveraged to support smarter program designs, with shared goals and objectives to help local civil society, human rights organizations, and independent media firmly establish their presence as part of efforts to combat the information disorder problem. As part of a push for collaboration, support multidisciplinary approaches. Some of the best examples in countering and preventing disinformation involve collaborative approaches that include data scientists, social scientists, and journalists.

10) **Measure the impact of your efforts to counter and prevent disinformation.** Monitoring and evaluation (M&E) are vital to helping USAID colleagues and others understand the relevance, effectiveness, efficiency, and impact of countering disinformation programs. Because there are so many unknowns in terms of what works and what does not in the field of countering disinformation, research and learning opportunities are important, and data collected should be shared with both USAID and its implementing partners. To support robust M&E, partner with local research firms, institutes or universities. There is a growing number of specialists with the right research skills and social science and computer science training to support your M&E programs in this space. Key skills include: social network analysis, content analysis and media monitoring, discourse and narrative analysis, and social media analysis. Before you begin new programs, make sure to capture baseline data, and include appropriate funds to also carry out an endline study. The "before and after" aspects of what you can capture through M&E are essential to program reporting.

ANNEX I: GLOSSARY OF TERMS

The field of disinformation has a particular lexicon: a stock of terms used to study and understand the broader field of information disorder and a particular vocabulary that is used by scholars and practitioners writing about the phenomenon. Definitions are collected from leading sources in the disinformation field.

Term	Definition
Algorithm	An algorithm is a fixed series of steps that a computer performs in order to solve a problem or complete a task. For instance, social media platforms use algorithms to compile the content that users see. These algorithms are designed to show users material that they will be interested in, based on each user's history of engagement on that platform. (Shorenstein Center, 2018)
Artificial Intelligence	Computer-based automated decision-making, inspired by human-like intelligence. Automated decisions might be directly implemented (e.g., in robotics) or suggested to a human decision-maker (e.g., product recommendations in online shopping). AI often incorporates machine learning (ML), in which predictions are based on patterns "learned" from existing data. (USAID, 2019) Also see *Machine learning*.
Astroturfing	Organized activity that is intended to create a false impression of a widespread, spontaneously arising, grassroots movement in support of or in opposition to something (such as a political policy) but that is initiated and controlled by a concealed group or organization (such as a corporation).
Backfire Effect	This effect is when beliefs are reinforced in the very attempt to debunk them.
Black box algorithms/ Black hat SEO (search engine optimization)	Describes aggressive and illicit strategies used to artificially increase a website's position within a search engine's results: for example, changing the content of a website after it has been ranked. These practices generally violate the given search engine's terms of service as they drive traffic to a website at the expense of the user's experience. (First Draft)
Bot	Bots are social media accounts that are operated entirely by computer programs and are designed to generate posts and/or engage with content on a particular platform. In disinformation campaigns, bots can be used to draw attention to misleading narratives, to hijack platforms' trending lists, and to create the illusion of public discussion and support. (Shorenstein Center, 2018) Also see *Sock puppet*.
Cheap Fakes	An audio or video manipulations created with cheaper, more accessible software (or none at all). They may be subtle, such as slowing down the speed at which a video is played, making it appear that the speaker's speech is slurred, or altering the background or an otherwise insignificant aspect of a picture.
Clickbait	Web content with misleading or sensationalist headlines that entices readers to click through to the full story, which is generally a disappointment. Clickbait's goal is usually to generate page views and advertising revenue. (Hootsuite, 2019)
Computational propaganda	Use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks.[231] (Woolley & Howard, 2017)
Content Farm	A website or company that creates low-quality content aimed at improving its search engine rankings. Also known as a content mill or factory, its main purpose is to maximize page views and revenue generated by advertising on those pages while minimizing the costs and time needed to create the content.[232]
Content Moderation	The process by which content is moderated on digital media platforms and users are warned/blocked, according to public terms of service agreements and platform

	"community standards." (Data & Society, 2019)
Coordinated Inauthentic Behavior	A term coined by Facebook in 2018 to describe the operation of running fake accounts and pages on an online social platform in order to influence discourse among users.
Cyber Troops	Government or political party actors tasked with the use of social media to manipulate public opinion online.[233]
Dangerous Speech	Any form of expression (speech, text or imaged) that can increase the risk that its audience will condone or participate in violence against members of another group.

Debunking	Publicly uncovering false information that is disseminated in order to influence or rather manipulate the whole of society or at least its major parts.[234]
Deep Fakes	Deepfakes are videos, images, and audio generated using artificial intelligence to synthetically render realistic depictions of speech and action.[235]
Digital Literacy	The ability to "access, manage, understand, integrate, communicate, evaluate and create information safely and appropriately through digital devices and networked technologies for participation in economic and social life. This may include competencies that are variously referred to as computer literacy, information and communication technology (ICT) literacy, information literacy, and media literacy." Digital literacy includes both hard skills related to the use of hardware or software and digital soft skills related to the use of digital media and information. (USAID, 2019)
Digital Media	Digital media includes the user-generated content and underlying software applications that make up internet-based communication tools, such as websites, mobile applications, news aggregators, social media platforms, search engines, and messaging/chat services. (USAID, 2019)
Digital Security	The practice of understanding one's digital footprint, identifying localized risks to information systems and taking reasonable steps to protect one's owned assets from loss or capture. (USAID, 2019)
Disinformation	Disinformation is false information that is deliberately created or disseminated with the express purpose to cause harm. Producers of disinformation typically have political, financial, psychological, or social motivations. (Shoranstein Center, 2018)
Doxing	Doxing, or doxxing, is the internet-based practice of researching and publicly broadcasting private or identifying information about an individual or organization. The methods employed to acquire this information include searching publicly available databases and social media websites, hacking, and social engineering.
Echo Chamber	An environment or social space where individuals are interacting with ideas and beliefs similar to their own. Existing ideas are thus reinforced, and they avoid being challenged by alternative ideas.[236]
Fact Checking	Fact-checking (in the context of information disorder) is the process of determining the truthfulness and accuracy of official, published information such as politicians' statements and news reports. (Shorenstein Center, 2018)
Filter Bubble	A situation that arises by virtue of an algorithmic filter, where internet and social media users interact with material that supports their own beliefs and ideas, and in turn algorithms continue to provide the user with similar content.[237]
Flooding	Flooding is spamming with unsolicited or misguided junk mail, chat, or social media messages such as advertising, brochures, or fake solicitations.
Gaslighting	A technique of deception and psychological manipulation practiced by a deceiver, or "gaslighter," on victims over an extended period. Its effect is to gradually undermine the victims' confidence in their own ability to distinguish truth from falsehood, right from wrong, or reality from appearance, thereby rendering them pathologically dependent on the gaslighter.[238]
Hate Speech	The use of speech to make direct attacks against an individual or a group of people

	based on a series of protected characteristics, such as race, ethnicity, nationality, religion, sex, sexual orientation, gender identity, and physical or mental ability. (USAID, 2019)
Inauthentic Actors	Individuals or organizations working to mislead others about who they are and what they are doing.[239]
Information Disorder	A condition in which truth and facts coexist in a milieu of misinformation and disinformation—conspiracy theories, lies, propaganda, and half-truths.
Internet Freedom	The U.S. Government conceptualizes internet freedom as the online exercise of human rights and fundamental freedoms regardless of frontiers or medium. The same rights that people have offline must also be protected online—in particular, freedom of expression, which is applicable regardless of frontiers and through any media of one's choice. (USAID, 2019)
Internet Governance	The development and application by governments, the private sector, and civil society, in their respective roles, of shared principles, norms, rules, decision-making

	procedures, and programs that shape the evolution and use of the internet. (United Nations, 2017)
Machine Learning	A set of methods for using computers to recognize patterns in data and make future predictions based on these patterns. Machine learning can be "supervised" or "unsupervised," depending on the level of human oversight. (USAID, 2019) Also see *Artificial intelligence.*
Malinformation	Deliberate publication of private information for personal or private interest, as well as the deliberate manipulation of genuine content. Note that these information are based on reality but are used and disseminated to cause harm. (Wardle & Derakhshan, 2017)
Manufactured Amplification	Boosting the reach or spread of information through artificial means.[240]
Media literacy	The ability to methodically consider and reflect on the meaning and source of a post or news article.
Meme	An idea or behavior that spreads from person to person throughout a culture by propagating rapidly and changing over time. The term is now used most frequently to describe captioned photos or GIFs that spread online; the most effective are humorous or critical of society. (Shorenstein Center, 2018)
Message Monitoring	Message monitoring analyzes the tropes, narratives, or specific messages that a bad actor is putting forward. In this way, it monitors platforms to look at the key messages that extremist or conspiracy groups are putting out to see if there are specific messages that they are repeating in talking points.
Microtargeting	Directing tailored advertisements, political messages, etc. at people based on detailed information about them (such as what they buy, watch, or respond to on a website); targeting small groups of people for highly specific advertisements or messages.[241]
Misinformation	Misinformation is information that is false, but not intended to cause harm. For example, individuals who do not know a piece of information is false may spread it on social media in an attempt to be helpful. (Shorenstein Center, 2018)
Natural Language Processing	The relationship between computers and human language content. It refers to speech analysis in both audible speech, as well as text of a language. NLP systems capture meaning from an input of words (sentences, paragraphs, pages, etc.).
Open-Source Intelligence (OSINT)	The multi-method approach of collecting and analyzing free, publicly available information and cross-referencing it against other public sources.[242]
Persona	The creation of a representative user based on available data and user interviews. Though the personal details of the persona may be fictional, the information used to create the user type is not. Personas can help bring a target audience to life. (Usability.gov, 2019)

Pink Slime Journalism	A low-cost way of distributing thousands of algorithmically generated news stories, often with political bias.
Platform Related	A social platform is a web-based technology that enables the development, deployment, and management of social media solutions and services. It provides the ability to create social media websites and services with complete social media network functionality. (Examples include Facebook, LinkedIn, Twitter, Snapchat, Pinterest, Instagram, WhatsApp, TikTok, and others.) Some disinformation is purposefully spread through algorithms that target specific audiences or inauthentic actors who gain a following in social media. Others are an unintended consequence of the way we share information or gain followers on social media platforms, such as the *Echo chamber* and the *Filter bubble*.
Prebunking	An offensive strategy that refers to anticipating what disinformation is likely to be repeated by politicians, pundits, and provocateurs during key events and having already prepared a response based on past fact checks.[243]
Propaganda	True or false information spread to persuade an audience but often has a political connotation and is often connected to information produced by governments.
Redirecting	Sending a user to a different site or reference that can serve to debunk or offer context to information presented in social media or on a website.
Satire	Satire is writing that uses literary devices such as ridicule and irony to criticize elements of society. Satire can become misinformation if audiences misinterpret it as fact. (Shorenstein Center, 2018)

Social Media Listening	A means of attaining interpersonal information and social intelligence from social media to understand how relationships are formed and influence the way we listen to and communicate with one another.[244]
Social Media Monitoring	The process of identifying and determining what is being said about an issue, an individual, or a group through different social and online channels. It is used also by businesses to protect and enhance the reputation of their brands and products. The method uses bots to crawl the internet and index messages based on a set of keywords and phrases.[245]
Sock Puppet	A sock puppet is an online account that uses a false identity designed specifically to deceive. Sock puppets are used on social platforms to inflate account's follower numbers and to spread or amplify false information to a mass audience. Often confused with a meat puppet, which is another actual individual using a false identity for the same purpose. (Shorenstein Center, 2018) Also see *Bot*.
Troll Farm	A troll farm is a group of individuals engaging in trolling or bot-like promotion of narratives in a coordinated fashion. (Shorenstein Center, 2018)
Trolling	The act of deliberately posting offensive or inflammatory content to an online community with the intent of provoking readers or disrupting conversation. The term "troll" is most often used to refer to any person harassing or insulting others online.
Verification	Verification is the process of determining the authenticity of information posted by unofficial sources online, particularly visual media. (Shorenstein Center, 2018)
Web Analytics	The measurement, collection, analysis and reporting of internet data for purposes of understanding and optimizing web usage. (Usability.gov, 2019)

ANNEX 2: TYPES OF MISINFORMATION & DISINFORMATION

The following examples were provided by First Draft News (https://firstdraftnews.org) to help illustrate and provide examples of the various types of misinformation and disinformation.

Fabricated content

https://www.vice.com/en_in/article/jged8b/the-first-use-of-deepfakes-in-indian-election-by-bjp
Indian politician uses deepfake to show himself giving a speech in a different language.

False Connection

https://www.youtube.com/watch?v=QiIhGzhW7o
The title of the video plays on anti-Chinese sentiment that is prevalent in Latin America to get people to click on the video and share. A textbook example of False Connection.

False context

https://twitter.com/AFPFactCheck/status/1221732075885101061
Video allegedly from Wuhan province where coronavirus originated is really from Indonesia.

Imposter content

https://twitter.com/elisethoma5/status/1215604086780743680?s=20
Fake screenshot that shows *Newsweek* article about Iran air strikes. Shows side-by-side comparisons of real/fake screens.

Satire or parody

https://factcheck.afp.com/australian-couple-quarantined-onboard-diamond-princess-cruise-reveal-wine-drone-delivery-story-was
The "news" about an Australian couple on cruise ship ordering wine via drone was debunked. The couple admitted that it was a joke post on Facebook for their friends.

Manipulated content

https://twitter.com/jamescracknell/status/1254395457033379843?s=21
Daily Mail edited a photo to make two people in a garden to appear closer than they really are.

Misleading content

https://www.bbc.com/news/blogs-trending-51020564
BBC Trending investigates cases of disinformation on Australian bushfires maps on social media.

ANNEX 3: EMERGING SOLUTIONS

Fact-Checking Initiatives

Africa Check is a nonprofit organization established in 2012 to promote accuracy in public debate and the media in Africa. Devised by the nonprofit media development arm of the international news agency AFP, Africa Check is an independent organization with offices in Johannesburg, Nairobi, Lagos, and Dakar. Africa Check produces reports in English and French, testing claims made by public figures, institutions and the media against the best available evidence. They have fact-checked more than 1,500 claims on topics from crime and race in South Africa to population numbers in Nigeria and fake health cures in various African countries.
Chequeado is the main project of the La Voz Pública Foundation. They are a nonpartisan and nonprofit digital medium dedicated to the verification of public discourse, the fight against disinformation, the promotion of access to information and the opening of data. Chequeado has been online since October 2010 and was the first site in Latin America dedicated to speech verification; it is among the top 10 fact-checking organizations in the world.
iRasKRIKavanje (Serbia) is a leading source for debunking disinformation in Serbia. With more than half a million monthly readers, it has been fact-checking media coverage of the COVID-19 pandemic, as well as covering the government's response to the crisis. Its posts have been regularly republished by the leading Serbian daily newspapers and portals.

PopUp Newsroom (global—some work has been in Mexico and Sweden) is a concept focused on collaboration between competing newsrooms in order to drive innovation, especially as it relates to the online disinformation sphere.[246] In Mexico, the organization helped launch *Verificado*, with more than 90 partners working to address disinformation around the 2018 elections.
TruthBuzz (global—with fellows and partners in Brazil, India, Indonesia, Nigeria, and U.S.) is an initiative sponsored by the International Center for Journalists to help reporters utilize compelling storytelling methods that improve the reach and impact of fact-checking and help audiences learn to identify the true from the false.[247] TruthBuzz fellows work together with a newsroom and receive training from **First Draft News**. One fellow's project supported Tempo.co, an Indonesian investigative reporting outlet, to reach a younger audience and begin introducing fact-checking approaches.[248]
Ukraine has developed some significant fact-checking initiatives. **VoxUkraine** uses a scientific analysis method to assess major economic and political processes and decisions in Ukraine. Among others, a key project by the Vox team is **VoxCheck**, a fact-checking service that identifies disinformation narratives being spread online. **TEXTY.org** is a Ukrainian nonprofit data analysis group that fact-checks and combats disinformation
What's Crap on WhatsApp (Africa) combats misinformation and disinformation on WhatsApp by utilizing the app itself to spread a fact-checking podcast. The show encourages users to send misinformation and disinformation to the organization's WhatsApp number and then the show producers create short podcast episodes that debunk the rumors and send them to their subscribers, who can easily forward them along to the groups in which the misinformation and disinformation spread in the first place.[249]

Examples of Elections-Focused Programming

ISFED's Fact-a-lyzer was created by the International Society for Fair Elections and Democracy, which was established for the 2018 presidential election in Georgia. Fact-a-lyzer is a pilot social media monitoring tool used for the election. The tool's software aggregated and monitored the Facebook content on public pages for political actors and groups, assigning tags to the posts. This data was then used to identify and analyze content trends.[250]
IFES's Social Media, Disinformation and Electoral Integrity examines the challenges disinformation presents to electoral integrity and the responses electoral management bodies and international nongovernmental organizations can take to mitigate threats.[251]
NDI's Disinformation and Electoral Integrity: A Guidance Document for NDI Elections Programs describes the National Democratic Institute's programmatic approaches to mitigating, exposing, and countering disinformation in the electoral context. The document stresses the importance of using open election data to deter disinformation and advocacy to counter it.[252]

Supporting Democracy's Guide for Civil Society on Monitoring Social Media During Elections provides an in-depth explanation of social media's role in elections, its impact on the electoral process, methodological approaches to monitoring it, data collection and analysis tool, and approaches for making an impact with social media monitoring.[253]

Examples of Coalitions

EUvsDisinfo is an EU-spearheaded project started in 2015 to "increase public awareness and understanding of the Kremlin's disinformation operations, and to help citizens in Europe and beyond develop resistance to digital disinformation and media manipulation."[254] EUvsDisinfo has hands in three major areas. First, it conducts data analytics of media spaces to identify and publicize disinformation campaigns started by the Kremlin or pro-Kremlin media outlets; the data are archived in an open-source, dedicated database. Second, the project publishes summaries, articles, and reports both within the research community and for wider general readership. Third, it offers training resources for governments, media outlets, universities, and civil society organizations.

The Atlantic Council is a "nonpartisan organization that galvanizes US leadership and engagement in the world, with allies and partners, to shape solutions to global challenges."[255] Among the numerous capacities through which Atlantic Council conducts international work, their recently published Democratic Defense Against Disinformation 2.0 provides a current snapshot of global disinformation campaigns at work, proposed solutions, progress made, and criticisms and recommendations made to the U.S. executive and judiciary branches regarding urgent steps needed in the path toward a healthier information ecosystem.

Resources for Legal and Policy Approaches

Poynter's **A Guide to Anti-misinformation Actions Around the World** has a regularly updated mapping by the Poynter Institute of both the positive and negative actions that governments have carried out to combat misinformation.[256]

UNESCO's **Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression** is unique in its global scale and comprehensiveness, but it is also highly action-oriented, with a suite of sector-specific actionable recommendations and a 23-point framework to test disinformation responses.

OxTEC's **A Report of Anti-Disinformation Initiatives** is a 2019 report by the Oxford Technology and Elections Commission (OxTEC) that analyzes fake news landscapes around the world, the governmental measures to address disinformation, and the geopolitical contexts facing each example. The report features examples from 19 countries in four continents.[257]

Examples of Counter-Disinformation Campaigns

A Counter-Disinformation System That Works details the formula **Debunk.eu** and its partners are using to counter disinformation in Lithuania. Using a combination of "geeks," or IT and AI experts developing algorithms to detect false claims; "elves," or volunteers who research and debunk false stories; and journalists, who publish finished stories about the debunked claims, Lithuanian anti-disinformation activists claim to have established a successful, fully integrated system.[258]

Russia's Disinformation Activities and Counter-Measures: **Lessons from Georgia** is a report from think-tank European Values about the main lessons learned from the fight against pro-Kremlin disinformation in Georgia.[259]

Examples of Local Journalism Support

Internews is aimed at building lasting change by ensuring people have access to quality, local information. To do so, Internews works with local partners to grow sustainable organizations and offers capacity-building programs for media professions, human rights activists, and information entrepreneurs.[260]

IREX promotes "vibrant information and media systems." IREX supports journalism and media organizations through trainings on reporting, media law, media safety, and digital security. IREX also provides additional support to consumers via media literacy programs, training citizen journalists, and diversifying and distributing television content.[261]

International Center for Journalists (ICFJ) seeks to build the expertise and storytelling skills of journalists around the world. ICFJ focuses on five key areas: news innovation, investigative reporting, global exchange programs, specialty journalism, and diversity promotion.[262]

Examples of Media Literacy

Digital Security Project seeks to provide data on the intersection of politics and social media. Its indicators cover topics including online censorship, polarization, misinformation campaigns, coordinated information operations and foreign influence in and monitoring of domestic politics. It uses the Varieties of Democracy (V-Dem) framework, also used by USAID in Journey to Self-Reliance (JSR) metrics, to assess various digital issues including misinformation.

IREX's Learn to Discern (L2D) is a worldwide media literacy training project for all ages that focuses on developing healthy information engagement habits and increasing the local demand for quality information. Its approach and curriculum are designed to meet the current needs of media consumers, adapting to the local context. L2D has been used in Ukraine, Serbia, Tunisia, Jordan, Indonesia, and the United States to address challenges stemming from misinformation, disinformation, propaganda, and influence campaigns. [263]

NewsWise is a free cross-curricular news literacy project for 9- to 11-year-old children across the United Kingdom, supported by the Guardian Foundation, National Literacy Trust, the PSHE Association, and Google. It features resources—including guides, webinars, and activities—for teachers and families.[264]

Examples of Public Awareness Campaigns

#ThinkB4UClick (Think Before You Click) is a campaign by #defyhatenow to raise awareness about the dangers of misinformation, fake news, and hate speech in South Sudan. It seeks to educate the public on these terms and explain how their individual actions can mitigate the issues to create safe online and offline spaces for healthy and informed discussions.

Elves vs. Trolls is an informal internet army of Lithuanians trying to counter what they describe as hate speech and pro-Russia propaganda.

The War on Pineapple, promoted by the U.S. Cybersecurity and Infrastructure Agency (CISA), uses the concept of pineapple on pizza to promote understanding of foreign interference in five steps: targeting divisive issues, moving accounts into place, amplifying and distorting the conversation, making the mainstream, and taking the conversation into the real world.

ANNEX 4: PASSIVE & ACTIVE DRIVERS OF DISINFORMATION

In *Demand for deceit: How the way we think drive disinformation*, Samuel Woolley and Katie Joseff compile a list of cognitive active and passive drivers that explain ways individuals can be manipulated into believing false content. These "cognitive drivers of consumption, acceptance, and sharing of disinformation" are broken into passive and active categories.

Passive Drivers

Belief Perseverance Effect	Continued influence of initial conclusions (sometimes based on false, novel information) on decision-making and individual beliefs.
Familiarity Effect	Information which is repeated or delivered in a manner consistent with past experience (for example, in a frequently heard accent) is often deemed more credible.
Misinformation Effect	False information suggested to individuals after the fact can influence their perception, especially as time passes and the memory weakens.
Priming	Shaping an individual's perceptions and behavior through exposure to subconscious stimuli.
Repeat Exposure	Individuals may respond more positively to stimuli that they have seen frequently than to stimuli they have seen only a few times; persists even when exposure is subliminal, and individuals are unaware that they have seen a stimulus.
Truth Bias	The default assumption that information is credible.
Virality and Heightened Emotion	Information which evokes fear, disgust, awe, anger, or anxiety may be much more likely to be spread by individuals over social media.

Active Drivers

Bandwagon Effect	The tendency of individuals to be more likely to adopt beliefs that they believe are common among others.
Confirmation Bias	Suggests that individuals seek out information that agrees with their preexisting beliefs.
Consensus Bias	The tendency to believe information that is perceived as consensus.
Disconfirmation Bias	Suggests that people actively reason against information which conflicts with preexisting beliefs.
Directionally Motivated Reasoning	The desire to reach a specific conclusion, and thus to lend more credibility to information favoring that conclusion.
In-group favoritism	The tendency to favor one's "in-group" (e.g., race, gender, sexual orientation, religious preference, partisan affiliation, geographic location, etc.) over one's out-group.
Preference Falsification	Occurs when individuals express preferences (e.g., favored politician or policy) in response to perceived societal pressures and do not communicate their true opinion.
Prior Attitude Effect	Suggests that people regard information that supports their beliefs ("pro-attitudinal information") as more legitimate than counter-attitudinal information (sometimes called the prior attitude effect).

Source: Woolley, S, & Joseff, K. Demand for deceit: How the way we think drives disinformation. Working paper. National Endowment for Democracy. https://www.ned.org/wp-content/uploads/2020/01/Demand-for-Deceit.pdf

ANNEX 5: QUICK RESOURCES FOR PLANNING A DISINFORMATION STRATEGY

1. Disinformation and Elections

- NDI Disinformation and Electoral Integrity
- IFES Social Media, Disinformation and Electoral Integrity
- Protecting Electoral Integrity in the Digital Age
- DRI Guide for Civil Society to Monitor Social Media During Elections

2. Disinformation Research Institutions

- Harvard Kennedy School/Misinformation Review
- First Draft News
- Stanford Cyber Policy Center
- Atlantic Council's Digital Disinformation Primer

3. Disinformation and Civil Society

- ComProp Navigator
- Interaction: disinformation toolkit for civil society

4. USG Resources for Analytics

- Global Engagement Center's DisinfoCloud
- Global Engagement Center's Counter-Disinformation Dispatches

5. USG Policy/Strategy

- National Security Strategy, December 2017 Pillar III: Preserve peace through strength"- Information Statecraft – "activate local network: local voices are most compelling and effective in ideological competitions"
- DoS/USAID Joint Strategic Plan FY 2018-2022 Strategic Objective 1.4: Increase capacity and strengthen resilience of our partners and allies to deter aggression, coercion and malign influence by state and non-state actors
- USAID Countering Malign Kremlin Influence Objective 2: Resist the manipulation of information
- National Defense Authorization Act 2016, 2017, 2020
- USG Organizations engaged on disinformation:
  - USG Organizations: Historical—The U.S. Information Agency
  - DoS/Global Engagement Center (GEC)
  - DoS/GPA Social Media Presence

- DoS/GPA Social Media Presence
- U.S. Agency for Global Media (USAGM)

ANNEX 6: SECTION-BY-SECTION RESOURCES

A. INTRODUCTION—SUPPLEMENTARY RESOURCES

The Wilson Center, the Carnegie Corporation of New York, and the University of Washington hosted a discussion on disinformation campaigns and potential ways to combat them. This video captures key aspects of the issues presented in the Primer. Panelists provided a historical overview of disinformation campaigns, assessed current patterns in disinformation, and discussed the potential challenges that lie ahead. Key concepts covered included: Disinformation Defined, Goal of Disinformation, Russian Disinformation, RT, Sputnik and Bots, and Identifying Disinformation.

https://www.c-span.org/video/?463432-1/panelists-discuss-combatting-disinformation-campaigns

ComProp Navigator is an online resource guide for CSOs to learn more about digital disinformation topics and address their concerns; it is curated by civil society practitioners and the Project on Computational Propaganda.[265]

https://navigator.oii.ox.ac.uk

B. PART TWO: UNDERSTANDING INFORMATION DISORDER—SUPPLEMENTARY RESOURCES

Dean Jackson for National Endowment for Democracy's International Forum Democratic Studies created three issue briefs to offer insights into various part of mis/disinformation:

- Issue Brief: Distinguishing Disinformation from Propaganda, Misinformation, and "Fake News" is helpful on defining disinformation.
- Issue Brief: How Disinformation Impacts Politics and Publics
- Issue Brief: The "Demand Side" of the Disinformation Crisis is helpful on how psychology and cognitive biases help perpetuate the reach and impact of disinformation.

For a look at how Finland has taken on mis/disinformation:

- Media Literacy in Finland Sets the Bar: Web Extra: Finnish Kids Got Education | Full Frontal on TBS
- Media literacy in Finland is the media literacy policy and the national media education policy document, published by the Ministry of Education and Culture in 2019.

For some insights into China:

- The Hoover Institution has a project called China's Global Sharp Power, which examines China's role in "digital authoritarianism," among other topics.
- Also read the exhaustive Combatting and Defeating Chinese Propaganda and

- Also read the exhaustive Combatting and Defeating Chinese Propaganda and Disinformation: A Case Study of Taiwan's 2020 Elections by Aaron Huang of the Belfer Center at Harvard Kennedy School.

To learn more on how algorithms amplify disinformation:

- What is an algorithm? in *The Economist*, August 30, 2017
- Misinformation has created a new world disorder: Our willingness to share content without thinking is exploited to spread disinformation by Claire Wardle in *Scientific American*, September 1, 2019.

- How misinformation spreads on social media—And what to do about it by Chris Meserole, Brookings Institution, May 9. 2018.

On how disinformation spreads across online platforms/applications:

- WhatsApp as a tool for fear and intimidation in Lebanon's protests, by Emily Lewis in Coda. November 12, 2019.
- Disinformation from China floods Taiwan's most popular messaging app, by Nithin Coca in Coda, October 7, 2020

C. Part Four: What Social Factors Contribute to Disinformation?

Resources for Media Monitoring:

- Bakamo.Social is a strategic social listening consultancy that uses technology and human understanding to find meaning in the millions of conversations that take place online every day. https://www.bakamosocial.com/
- East Stratcom Task Force (EU) was set up to address Russia's ongoing disinformation campaigns. In March 2015, the European Council tasked the High Representative in cooperation with EU institutions and Member States to submit an action plan on strategic communication. https://eeas.europa.eu/headquarters/headquarters-homepage/2116/-questions-and-answers-about-the-east-stratcom-task-force_en
- KremlinWatch is a strategic program of the European Values Center for Security Policy, which aims to expose and confront instruments of Russian influence and disinformation operations focused against Western democracies. https://www.kremlinwatch.eu/
- Memo 98 is a Slovakia-based specialist media monitoring organization, with extensive experience of delivery media analyses on behalf of international institutions as well as technical assistance to civil society groups. It offers specialized media monitoring services in the area of disinformation.
- Moonshot Consulting seeks to find audiences vulnerable to violent extremist and false messaging, works to better understand them, and then builds an evidence base to deliver campaigns and interventions to make information safer. http://moonshotcve.com/work/
- Researchers from the University of Notre Dame are using artificial intelligence to develop an early warning system that will identify manipulated images, deepfake videos, and disinformation online. The project is an effort to combat the rise of coordinated social media campaigns to incite violence, sow discord, and threaten the integrity of democratic

elections. https://news.nd.edu/news/researchers-develop-early-warning-system-to-fight-disinformation-online/

Examples of open-source intelligence (OSINT):

- "How to Conduct an Open-Source Investigation, According to the Founder of Bellingcat" is an article about Eliot Higgins, founder of the open-source investigation website Bellingcat, and his workshops to teach journalists, NGOs, government agencies, and other interested parties how to use OSINT for their investigations. Bellingcat provides an online investigation toolkit that is updated regularly and provides open-source and free software.[266]
- First Draft Basic Toolkit also provides links to open-source and free software of use to newsrooms for newsgathering, verification, and responsible reporting.
- Atlantic Council's Digital Forensic Research Lab (DFRLab) uses open-source research to expose and explain disinformation; the DFRLab seeks to build "the world's leading

hub of digital forensics analysts (#DigitalSherlocks)," promoting objective truth, protecting democratic institutions and norms, and forging greater digital resilience worldwide.[267]

- "Civil Society Tracks Trolls and Fakes, Prompts Facebook Action in Moldova" is an article describing Trolless, a platform that enabled Moldovan users to report fake profiles, troll accounts, and suspicious activity and material. Once reported, the Trolless team would investigate and publish their findings, providing verification or more information about the suspected accounts and content.[268]

D. Part Five: What Are Some Anticipated Challenges?

To learn more about some of the technologies mentioned in this section:

- The Hamilton 2.0 dashboard, a project of the Alliance for Securing Democracy at the German Marshall Fund of the United States, provides a summary analysis of the narratives and topics promoted by Russian, Chinese, and Iranian government officials and state-funded media on Twitter, YouTube, state-sponsored news websites, and via official diplomatic statements at the United Nations. https://securingdemocracy.gmfus.org/hamilton-dashboard/
- Artificial Intelligence (AI)-Generated Propaganda lab OpenAI has already released a beta version of GPT-3, a long-form text generator that works by taking text input and predicting what should follow. https://openai.com/blog/openai-api/

E. Part Six: What Are Some Emerging Solutions for Disinformation?

On debunking and discrediting:

- The Global Engagement Center (GEC) at the U.S. Department of State recommends a combined debunking and discrediting approach, which is explained in GEC Counter-Disinformation Dispatches #2: Three Ways to Counter Disinformation and GEC Counter-Disinformation Dispatches #4: What Works in Debunking.

ANNEX 7: WHAT TO READ & WATCH

Select key books to help orient your knowledge and understanding of disinformation studies

- *Algorithms of Oppression: How Search Engines Reinforce Racism* by Safiya Umoja Noble
- *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy* by Siva Vaidhyanathan
- Communication Power by Manuel Castells
- *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* by Samuel C. Woolley and Philip N. Howard (Editors)
- *How Propaganda Works* by Jason Stanley
- *How to Lose the Information War: Russia, Fake News, and the Future of Conflict* by Nina Jankowicz
- *Lie Machines: How to Save Democracy From Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives* by Philip Howard
- *Mind Over Media: Propaganda Education for a Digital Age* by Renee Hobbs and Douglas Rushkoff
- *Nothing Is True and Everything Is Possible: The Surreal Heart of the New Russia The Surreal Heart of the New Russia* by Peter Pomerantsev

- *This Is Not Propaganda: Adventures in the War Against Reality Adventures in the War Against Reality* by Peter Pomerantsev
- *Twitter and Tear Gas: The Power and Fragility of Networked Protest* by Zeynep Tufekci

Key podcasts of interest

- All Things Policy, Episode 354: *Ordering the Information Disorder:* https://player.fm/series/all-things-policy/ep-354-ordering-the-information-disorder
- CSIS, *Confronting the Problem of Fake News in Africa:* https://www.csis.org/events/confronting-problem-fake-news-africa
- Democracy in Danger; produced by the Deliberative Media Lab with support from the UVA Democracy Initiative and the College of Arts and Sciences: https://medialab.virginia.edu/democracyindanger
- Demystifying Media at the University of Oregon, Episode 39: *Fighting a New Era of Disinformation with Claire Wardle:* https://www.stitcher.com/podcast/demystifying-media-podcast
- *The Disinformation Age:* https://www.mikehind.co.uk
- Interpret; *Health Misinformation & Polio in Pakistan with Carlotta Dotto:* https://www.buzzsprout.com/739217/4387205-health-misinformation-polio-in-pakistan-with-carlotta-dotto
- NED's Power 3.0; blog posts and podcast episodes related to media and technology: https://www.power3point0.org/category/media-information-technology/

- New York Public Radio, On the Media; *Ukraine's Remedy for Fake News: News About Fake News:* https://www.wnycstudios.org/podcasts/otm/segments/ukraine-remedy-fake-news-more-news
- The Nordic Co-operation, The Foreign Desk; *The Threats to Democracy:*
- https://www.norden.org/en/information/foreign-desk-threats-democracy
- NPR, Rough Translation; *Ukraine vs. Fake News:* https://www.npr.org/2017/10/17/544458898/ukraine-vs-fake-newsPodcast Brunch Club: Disinformation and Fake News; January 2020 Listening List: https://podcastbrunchclub.com/fakenews/
- The Poynter Institute; *Is fact-checking the antidote to misinformation?* https://www.spreaker.com/user/newsu/misinformed-episode-3?utm_medium=widget&utm_source=user%3A10628419&utm_term=episode_titleRecode Decode, Episode 524; *Phil Howard and Emily Bell: Disinformation in 2020, from "Plandemic" to Bill Gates to "Obamagate": while* US-focused, this podcast contains very useful information and explanations that will be helpful to teaching and understanding the disinformation challenge

Documentaries and Useful YouTube Videos

- *After Truth: Disinformation and the Cost of Fake News:*

- Al Jazeera—*India: Fake News and Agitprop:* https://www.youtube.com/watch?v=S0LgwL6rMPk&list=PLzGHKb8i9vTwQ4uKHdPDjqhnAap_h1mdh&index=39
- Al Jazeera—*How Fake News Could Shape Kenya's Election:* https://www.youtube.com/watch?v=Pz65C9fQpdM&list=PLzGHKb8i9vTwQ4uKHdPDjqhnAap_h1mdh&index=13
- Al Jazeera—Inside Story: *Should Social Media Be Regulated?* https://www.youtube.com/watch?v=-QaXVetsuIg&list=PLzGHKb8i9vTwQ4uKHdPDjqhnAap_h1mdh&index=56
- *Fact or Friction: Reporting on Hong Kong's Protests:* https://www.youtube.com/watch?v=blUwA7NkITU&list=PLzGHKb8i9vTwQ4uKHdPDjqhnAap_h1mdh&index=53
- MIT's *Moon Disaster:* https://moondisaster.org
- New York Times' *Operation Infektion:* https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html
- TED—*How Can We Protect in the Age of Misinformation: Sinan Aral: https://www.youtube.com/watch?v=-7ORAKULeI4*
- TED Global—*How to Seek Truth in an Era of Fake News: Christiane Amanpour:* https://www.ted.com/talks/christiane_amanpour_how_to_seek_truth_in_the_era_of_fake_news?referrer=playlist-the_value_of_skepticism
- *A Thousand Cuts:* https://www.athousandcuts.film
- *Trust Me* Documentary: https://www.trustmedocumentary.com/?gclid=CjwKCAjww5r8BRB6EiwArcckC9U5UTNVNdfKHjV_GGfFGXTu8DQNjLqltYKIRYKGc-TJwtjpPmyX8RoCSCoQAvD_BwE
- *Washington* Post's *Fakeout* Series: https://www.youtube.com/watch?v=KwkoFkA2UoI

[1] Vosoughi, S, et al. (2018). The spread of false information online. *Science* (359), 1146-1151. https://science.sciencemag.org/content/sci/359/6380/1146.full.pdf

[2] See: Prevency. *What is disinformation?* prevency.com/en/what-is-disinformation/

[3] Gunitsky, S. (2020, April 21). *Democracies can't blame Putin for their disinformation problem.* Foreign Policy. Opinion. https://foreignpolicy.com/2020/04/21/democracies-disinformation-russia-china-homegrown/

[4] Oxford Internet Institute Report. (2019). *Use of social media to manipulate public opinion now a global problem.* https://www.oii.ox.ac.uk/news/releases/use-of-social-media-to-manipulate-public-opinion-now-a-global-problem-says-new-report/

[5] Gunitsky, S. (2020, April 21). *Democracies can't blame Putin for their disinformation problem.* Foreign Policy. Opinion. https://foreignpolicy.com/2020/04/21/democracies-disinformation-russia-china-homegrown/

[6] See the EU's 2019 concept note "How to spot when news is fake": https://epthinktank.eu/2018/04/24/online-disinformation-and-the-eus-response/how-to-spot-when-news-is-fake-blog/

[7] See: Storyful Intelligence. (2018. September 24). Misinformation and Disinformation. White paper. Storyful. https://storyful.com/thought-leadership/misinformation-and-disinformation/

[8] Pennycook, G., & Rand, D. G. (2018). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* (188): 39-50. https://doi.org/10.1016/j.cognition.2018.06.011

[9] Johnston, J. (2020). *Disinformation poses 'existential threat' to democracy, parliamentary committee warns.* PublicTechnology.net. https://www.publictechnology.net/articles/news/disinformation-poses-%E2%80%98existential-threat%E2%80%99-democracy-parliamentary-committee-warns

[10] Vaidhyanathan, S. (2018). *Antisocial media: How Facebook disconnects us and undermines democracy.* New

York: Oxford University Press.

[11] "Autocratic rulers … benefit from the distrust, cynicism, and social atomization produced by disinformation, precisely because it inhibits political engagement and paralyzes organized social movements. Social media echo chambers, deepfakes, and automated surveillance have all served to lower the costs of autocracy while undermining the benefits of open democratic deliberation. Far from being overwhelmed by free information, dictators are increasingly learning to take advantage of it." From: Gunitsky, S. (2020, April 21). *Democracies can't blame Putin for their disinformation problem.* Foreign Policy. Opinion. https://foreignpolicy.com/2020/04/21/democracies-disinformation-russia-china-homegrown/

[12] See, for example: Fremis, A. G. (2020, November 1). *Combating disinformation: Policy analysis and recommendations for the 21st century information environment.* Atlantic Forum. https://atlantic-forum.com/content/combating-disinformation-policy-analysis-and-recommendations-21st-century-information

[13] Policy Department for Citizens' Rights and Constitutional Affairs. (2019). *Disinformation and propaganda: Impact on the functioning of the rule of law in the EU and its Member States.*

[14] *Doxing* and *trolling* are forms of online harassment. See: PEN America's Online Harassment Field Manual for a very helpful overview of the key terms and breakdown of different categories related to cyber-harassment or cyber-abuse: https://onlineharassmentfieldmanual.pen.org/defining-online-harassment-a-glossary-of-terms/

[15] *Echo chamber* is where a group chooses to preferentially connect with each other, to the exclusion of outsiders.
Filter bubble is where a group chooses to preferentially communicate with each other, to the exclusion of outsiders. See Annex I for a glossary of useful disinformation terms.

[16] Clare Wardle is the cofounder of First Draft, a nonprofit organization formed in 2016 to protect communities from harmful misinformation. Wardle is considered a leading expert on information disorder. See: https://firstdraftnews.org

[17] Wardle, C. (2020). *Training: Understand the landscape of information disorder.* First Draft. https://firstdraftnews.org/training/information-disorder/

[18] Habgood-Coote, J. (2018). *The term 'fake news' is doing great harm.* The Conversation. https://theconversation.com/the-term-fake-news-is-doing-great-harm-100406

[19] Dema, T. (2017). *Media literacy vital to prebunk and debunk fake news.* Kuensel. lkyspp.nus.edu.sg/docs/default-source/ips/kuensel_media-literacy-vital-to-prebunk-and-debunk-fake-news_190817-pdf.pdf?sfvrsn=c41b910b_0

[20] National Endowment for Democracy. (2020). *From democratic regression to 'third reverse wave.'* Democracy Digest. https://www.demdigest.org/from-democratic-regression-to-third-reverse-wave/

[21] Freedom House. (2019). *Countries and territories.* https://freedomhouse.org/countries/freedom-world/scores

[22] Freedom House. (2012). https://freedomhouse.org/sites/default/files/2020-02/FIW_2010_Overview_Essay.pdf

[23] The topic of zero-rated content such as Facebook's Free Basics is of considerable debate in the internet freedom and digital rights community and increasingly for independent media support. According to a 2017 report by Berkman Klein Center, "Zero rating, which allows users to access select Internet services and content without incurring mobile data charges, is not a new concept. But it has become an object of debate as mobile carriers and major app providers have used it in the developing world to attract customers, with the goal of increasing Internet access and adoption." For the full report, see: https://cyber.harvard.edu/publications/2017/10/zerorating. A significant number of USAID partner countries receive Internet access through the zero-rating scheme: https://en.wikipedia.org/wiki/Zero-rating

[24] Reuters. (2020, August 11). *UN investigator says Facebook has not shared 'evidence' of Myanmar crime.* Malay Mail. https://www.malaymail.com/news/world/2020/08/11/un-investigator-says-facebook-has-not-shared-evidence-of-myanmar-crime/1892905

[25] Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook J. (2012, December). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest 13*(3):106-31. https://journals.sagepub.com/doi/full/10.1177/1529100612451018

[26] 2020 Edelman Trust Barometer. (2020, January 19). https://www.edelman.com/trustbarometer

[27] Associated Press. (2020, February 7). *Cyborgs, trolls and bots: A guide to online misinformation.* Snopes. https://www.snopes.com/ap/2020/02/07/cyborgs-trolls-and-bots-a-guide-to-online-misinformation/

[28] See: Bennett, L, & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication.*

[29] Edelman Trust Barometer 2020 Global Report. https://cdn2.hubspot.net/hubfs/440941/Trust%20Barometer%202020/2020%20Edelman%20Trust%20Barometer%20Global%20Report.pdf?utm_campaign=Global:%20Trust%20Barometer%202020&utm_source=Website

[30] Proquest. How to identify fake news in 10 steps. Worksheet. https://blogs.proquest.com/wp-content/uploads/2017/01/Fake-News1.pdf

[31] Statista. (2020, January.) *Global internet penetration rate as of January 2020, by region.* https://www.statista.com/statistics/269329/penetration-rate-of-the-internet-by-region/

[32] Silver, L. (2019). *Smartphone ownership is growing rapidly around the world, but not always equally.* Pew Research Center's Global Attitudes Project. https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/

[33] Buchholz, K. (2020). *Where do people spend the most time on social media.* Statista. https://www.statista.com/chart/18983/time-spent-on-social-media/

[34] Lamb, K. (2019). *Philippine tops world internet usage index with an average of 10 hours a day*. The Guardian. https://www.theguardian.com/technology/2019/feb/01/world-internet-usage-index-philippines-10-hours-a-day
[35] Smith, K. *60 incredible and interesting Twitter stats and statistics*. Brandwatch. https://www.brandwatch.com/blog/twitter-stats-and-statistics/
[36] Statista. (2020, July). *Leading countries based on Facebook audience size as of July 2020*. https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/
[37] Statista. (2020, June). *Share of adults who use social media as a sources of news in selected countries*. Statista. https://www.statista.com/statistics/718019/social-media-news-source/
[38] Farrell, M. (2019). *The Internet must be more than Facebook*. OneZero. https://onezero.medium.com/the-internet-must-be-more-than-facebook-4ba1d86403fb
[39] Farrell, M. (2019). *The Internet must be more than Facebook*. OneZero. https://onezero.medium.com/the-internet-must-be-more-than-facebook-4ba1d86403fb
[40] Moon, M. (2017, June 25). *WhatsApp is becoming a top news source in some countries*. Engadget. https://www.engadget.com/2017-06-25-whatsapp-news-source-reuters-study.html
[41] Sahir. (2019, February 11). *WhatsApp usage, revenue, market share and other statistics (2019)*. Digital Information World. https://www.digitalinformationworld.com/2019/02/whatsapp-facts-stats.html
[42] Silver, L. (2019). *Smartphone ownership is growing rapidly around the world, but not always equally*. Pew Research Center's Global Attitudes Project. https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/
[43] Lardieri, A. (2019). Older people more susceptible to take news, more likely to share it. https://www.usnews.com/news/politics/articles/2019-01-09/study-older-people-are-more-susceptible-to-fake-news-more-likely-to-share-it
[44] Loos, E. & Nijenhuis, J. (2020). Consuming Fake News: A Matter of Age? The perception of political fake news stories in Facebook ads. https://www.researchgate.net/publication/336944922_Consuming_Fake_News_A_Matter_of_Age_The_perception_of_political_fake_news_stories_in_Facebook_ads
[45] Kight, S. W. (2020). *Gen Z is eroding the power of misinformation*. Axios. https://www.axios.com/gen-z-is-eroding-the-power-of-misinformation-5940e3cd-e3d0-44a1-b66c-93be45fe1d2c.html
[46] Kovacs, K. (2020). *Gen Z has a misinformation problem*. Digital Content Next. https://digitalcontentnext.org/blog/2020/06/04/gen-z-has-a-misinformation-problem
[47] Mackintosh, Eliza, for CNN. Finland is winning the war on fake news. What it's learned may be crucial to Western democracy. https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/
[48] Brashier, N. M., et al. (2017). Competing cues: Older adults rely on knowledge in the face of fluency. *Psychology and aging* (32)4. 331-337. doi.10.1037/pag0000156
[49] Jacoby, L. L., & Rhodes, M. G. (2006). False remembering in the aged. *Current Directions in Psychological Science* (15)2: 49-53. https://doi.org/10.1111/j.0963-7214.2006.00405.x
[50] Silver, L. (2019). *Smartphone ownership is growing rapidly around the world, but not always equally*. Pew Research Center's Global Attitudes Project. https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/
[51] Vered, E. (2020). *Middle Eastern journalists targeted by misogynistic smear campaigns*. Newsroom. International Press Institute. https://ipi.media/middle-eastern-journalists-targeted-by-misogynistic-smear-campaigns/
[52] Vered, E. (2020). *Middle Eastern journalists targeted by misogynistic smear campaigns*. Newsroom. International Press Institute. https://ipi.media/middle-eastern-journalists-targeted-by-misogynistic-smear-campaigns/
[53] Gendered Disinformation, Fake News, and Women in Politics: https://www.cfr.org/blog/gendered-disinformation-fake-news-and-women-politics
[54] Ayyub, R. (2018). *I was the victim of a deepfake porn plot intended to silence me: Rana Ayyub*. HuffPost India. https://www.huffingtonpost.in/rana-ayyub/deepfake-porn_a_23595592/
[55] Ayyub, R. (2018). *I was the victim of a deepfake porn plot intended to silence me: Rana Ayyub*. HuffPost India. https://www.huffingtonpost.in/rana-ayyub/deepfake-porn_a_23595592/
[56] Reporters Without Borders. (2018). *Women's rights: Forbidden subject*. Reporters Without Borders. https://rsf.org/sites/default/files/womens_rights-forbidden_subject.pdf
[57] Posetti, J., Harrison, J., & Waisbord, S. *Online attacks on women journalists leading to 'real world' violence, new research shows*. ICFJ. https://www.icfj.org/news/online-attacks-women-journalists-leading-real-world-violence-new-research-shows
[58] Macavaney, S, et al. (2019). Hate speech detection: Challenges and solutions. *Plos One* (14)8. https://doi.org/10.1371/journal.pone.0221152

[59] Macavaney, S, et al. (2019). Hate speech detection: Challenges and solutions. *Plos One* (14)8. https://doi.org/10.1371/journal.pone.0221152
[60] PeaceTech Lab. *Combating hate speech: Identifying, monitoring and combating hate speech on social media*. https://www.peacetechlab.org/hate-speech.
[61] Williams, M. L., Burnap, P., Javed, A., Liu, H., & Ozalp, S. (2020). Hate in the machine: Anti-Black and anti-Muslim social media posts as predictors of offline racially and religiously aggravated crime. *British Journal of*

Criminology. https://academic.oup.com/bjc/article/60/1/93/5537159
[62] Dangerous Speech Project. https://dangerousspeech.org/
[63] Dangerous Speech Project. https://dangerousspeech.org/
[64] Stricklin, K. (2020). *Why does Russia use disinformation?* Lawfare. https://www.lawfareblog.com/why-does-russia-use-disinformation
[65] Nimmo, B. (2015, May 19). *Anatomy of an info-war: How Russia's propaganda machine works, and how to counter it.* StopFake.org. https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/
[66] Ellick, A.B., & Westbrook, A. (2018). *Operation Infektion: A three-part video series on Russian disinformation.* New York Times. https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html
[67] Young, C.W. (1987). *Congressional Record: Soviet active measures in the United States—An updated report by the FBI.* New York Times. https://www.cia.gov/library/readingroom/document/cia-rdp11m01338r000400470089-2
[68] Ellick, AB. & Westbrook, A. (2018). *Operation Infektion: A three-part video series on Russian disinformation.* New York Times. https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html
[69] Bugayova, N, & Barros, G. (2020). *The Kremlin's expanding media conglomerate.* Institute for the Study of War. http://www.understandingwar.org/backgrounder/kremlin%E2%80%99s-expanding-media-conglomerate
[70] Wardle, C., & Derakhshan, H. (2017). Information disorder: Towards an interdisciplinary framework for research and policy making. Council of Europe report. https://rm.coe.int/information-disorder-report-version-august-2018/16808c9c77
[71] Nimmo, B. (2015, May 19). Anatomy of an info-war: How Russia's propaganda machine works, and how to counter it. StopFake.org. https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/
[72] The White Helmets are a volunteer organization that operates in parts of opposition-controlled Syria and in Turkey. Formed in 2014 during the Syrian Civil War, the majority of the volunteers' activity in Syria consists of medical evacuation, urban search and rescue in response to bombing, evacuation of civilians from danger areas, and essential service delivery.
[73] Chulov, M. (2020, October 27). *How Syria's disinformation wars destroyed the co-founder of the White Helmets. Guardian.* https://www.theguardian.com/news/2020/oct/27/syria-disinformation-war-white-helmets-mayday-rescue-james-le-mesurier
[74] Ellick, A. B., & Westbrook, A. (2018). *Operation Infektion: A three-part video series on Russian disinformation.* New York Times. https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html
[75] GEC. (August 2020). Pillars of Russia's disinformation and propaganda ecosystem. Special report. https://www.state.gov/wp-content/uploads/2020/08/Pillars-of-Russia%E2%80%99s-Disinformation-and-Propaganda-Ecosystem_08-04-20.pdf
[76] Nimmo, B. (2015, May 19). *Anatomy of an info-war: How Russia's propaganda machine works, and how to counter it.* StopFake.org. https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/
[77] Ellick, A. B., & Westbrook, A. (2018). *Operation Infektion: A three-part video series on Russian disinformation.* New York Times. https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html
[78] Statement of Lea Gabrielle, Special Envoy & Coordinator for the Global Engagement Center, U.S. Department of State, Before the Senate Foreign Relations Subcommittee on State Department and USAID Management, International Operations, and Bilateral International Development, Thursday, March 5, 2020, https://www.foreign.senate.gov/imo/media/doc/030520_Gabrielle_Testimony.pdf
[79] Statement of Lea Gabrielle, Special Envoy & Coordinator for the Global Engagement Center, U.S. Department of State, Before the Senate Foreign Relations Subcommittee on State Department and USAID Management, International Operations, and Bilateral International Development, Thursday, March 5, 2020. https://www.foreign.senate.gov/imo/media/doc/030520_Gabrielle_Testimony.pdf
[80] See: Swan, B. W. (2020, April 21). *State report: Russian, Chinese and Iranian disinformation narratives echo one another.* Politico. https://www.politico.com/news/2020/04/21/russia-china-iran-disinformation-coronavirus-state-department-193107
[81] Vilmer, J. J., & Charon, P. (2020, January 21). *Russia as a hurricane, China as climate change: Different ways of information warfare.* War on the Rocks. https://warontherocks.com/2020/01/russia-as-a-hurricane-china-as-climate-change-different-ways-of-information-warfare/
[82] Vilmer, J. J., & Charon, P. (2020, January 21). *Russia as a hurricane, China as climate change: Different ways of information warfare.* War on the Rocks. https://warontherocks.com/2020/01/russia-as-a-hurricane-china-as-climate-change-different-ways-of-information-warfare/
[83] Vilmer, J. J., & Charon, P. (2020, January 21). *Russia as a hurricane, China as climate change: Different ways of information warfare.* War on the Rocks. https://warontherocks.com/2020/01/russia-as-a-hurricane-china-as-climate-change-different-ways-of-information-warfare/
[84] Vilmer, J. J., & Charon, P. (2020, January 21). *Russia as a hurricane, China as climate change: Different ways of information warfare.* War on the Rocks. https://warontherocks.com/2020/01/russia-as-a-hurricane-china-as-climate-change-different-ways-of-information-warfare/